Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
AIR FORCE PERSONNEL AND TRAINING RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Clotus J. Burke, *Indiana University*
Robert M. Gagné, *AF Personnel and Training Research Center*
W. R. Garner, *The Johns Hopkins University*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Arthur L. Irion, *Tulane University*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

The Summation of Generalized Reactive Tendencies: E. A. Bilodeau, J. S. Brown, and J. J. Meetman 293

Experimental Extinction and Drive During Extinction in a Discrimination Habit: J. R. Cautela 299

Delay of Reward and Performance of an Instrumental Response: G. S. Harker 303

Inferential Behavior in Preschool Children: H. H. Kendler and T. S. Kendler 311

Two-Choice Behavior of Paradise Fish: R. R. Bush and T. R. Wilson 315

Studies in Incidental Learning: III. Interserial Interference: L. Postman and P. A. Adams 323

Studies in Incidental Learning: IV. The Interaction of Orienting Tasks and Stimulus Materials: L. Postman and P. A. Adams 329

Studies in Incidental Learning: V. Recall for Order and Associative Clustering: L. Postman, P. A. Adams, and A. M. Bohm 334

Perceptual Thresholds of Curvilinearity and Angularity as Functions of Line Length: L. Della Valle, T. G. Andrews, and S. Ross 343

Sugar as a Reward for Hungry and Nonhungry Rats: M. Smith and G. C. Kinney 348

Reactive Inhibition as a Function of Same-Hand and Opposite-Hand Intertrial Activity: L. E. Albright, C. R. Bonsberg, and M. H. Marx 353

Retention of Concepts as a Function of the Degree of Original and Interpolated Learning: J. Richardson 358

American Psychological Association

Vol. 51 No. 5 May 1956

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1956 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, Box 1557, Lackland Air Force Base, San Antonio, Texas. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail.

To save printing costs supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S's*, and *Ss'*; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 15th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1956 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 51, NO. 5 MAY, 1956

THE SUMMATION OF GENERALIZED REACTIVE TENDENCIES

EDWARD A. BILODEAU, JUDSON S. BROWN, AND JOHN J. MERYMAN[1]

State University of Iowa

The most explicit theoretical treatment of stimulus generalization is probably that presented by Hull in his paper on stimulus equivalence (**4**). Using, as his point of departure, Hovland's (**3**) data on the generalization of the GSR to auditory stimuli of various frequencies, Hull develops some 29 theorems concerning generalization gradients and their interactions. One of his most interesting theorems deals with the summation of overlapping generalization gradients. According to this theorem, if a response has been conditioned to two distinct points on the same stimulus continuum, and if these points are not too remote from one another, the generalized excitatory tendencies to respond to stimuli intermediate between the points should summate to yield a stronger reactive tendency than that of either generalized tendency alone. Moreover, if the generalized tendencies from either of the points extend beyond the other, the overlapping tendencies should also summate according to the same rule.

Although this theorem seems to have been stated with sufficient clarity to make its experimental study feasible, apparently no one has done so to date. Since Hull's theory was presented within the framework of classical conditioning, failures to test its theorems experimentally may be a consequence of the difficulties that often attend the setting up of stable conditioned reactions in both human and animal subjects.

In a previous paper by Brown, Bilodeau, and Baron (**1**) a method of studying the empirical phenomena of generalization was proposed that yielded results analogous to those obtained with the methods of classical conditioning while circumventing certain of the difficulties inherent in the latter. This method, which was quite similar to that used by Gibson (**2**), involved voluntary manual movements rather than classical conditioned reactions. Briefly, *S*s were instructed to respond as quickly as possible to the lighting of the middle one of seven lamps (arranged in a horizontal row) by lifting their fingers

[1] The first author is now at Tulane University; the third author is now at Lake Forest College; the second author is primarily responsible for the preparation of the manuscript.

from a telegraph key. They were told not to respond to the lighting of any of the other six lamps. Test presentations of each of the forbidden lamps revealed that the nearer a lamp was to the middle one, the more likely it was to elicit prohibited, or what may be called *generalized*, responses. Plots of response frequency against lamp position yielded negatively accelerated, regular gradients on each side of the central stimulus, with essentially identical curves being obtained in three independent repetitions of the experiment. In discussing the results, the authors suggested that Hull's concept of generalized habit strength might be extended to the interpretation of bidirectional gradients in the frequency of these manual responses.

Since the method proved practical and economical for the investigation of simple generalization gradients along a visual-spatial dimension, it seemed worth while to apply it to the study of more complex phenomena such as that predicted by Hull's summation theorem. This paper reports such an investigation.

In broad outline, the present experiment involved a procedure in which *Ss* were instructed to make a manual response to the lighting of either one of two spatially separated lamps in a horizontally arranged row of seven. They were also instructed to refrain from responding to any of the remaining five lamps. From a series of trials to both negative and positive lamps it was possible to determine whether frequency of response to negative lamps situated between or beyond the two positive ones would vary in a manner coordinate with the hypothesis that overlapping generalized tendencies summate.

Method

Apparatus.—The apparatus was essentially the same as that described in the previous paper (1) where a more detailed description may be found. The principal component was a black, curved stimulus-lamp board on which seven 7.5-w. lamps were placed in a horizontal row at eye level, separated by 8° of visual angle. Each lamp was 5 ft. from the eyes of the seated *S* who faced directly toward the middle lamp. For convenience in referring to the lamps, they are denoted hereafter by the numbers 1 through 7, from left to right, as seen by *S*. A small platform at table height supported *S*'s hand and his reaction key. By means of suitable switches and timing equipment, *E*, who was seated behind the lamp board, could turn on any of the lamps and control the foreperiod and intertrial intervals. Reaction times were observed by *E* from a Standard Electric Timer, though they were used solely to keep a constant check on *S*'s attentiveness to the task. The *S* was urged to react more quickly whenever two or three successive reaction times exceeded .5 sec.

Two minor modifications of the original apparatus were made. A small white fixation cross was substituted for the original green-jeweled pilot lamp located directly above the center lamp. At the start of the foreperiod, a warning buzz was presented through a pair of headphones to indicate that a lamp was about to light. The green-jeweled lamp had formerly served both as ready signal and as fixation point.

Procedure.—The basic design involved two experimental groups of *Ss*. One of these, Group 2–6, was instructed to react as quickly as possible to the lighting of either Lamp 2 or 6; the other, Group 3–5, was told to react when either 3 or 5 was lighted. Both groups were instructed not to react to lamps other than those specifically designated. In all other respects the groups received identical treatments.

As in the original experiment, *Ss* were told that the purpose of the study was to measure reaction time and that their task was to lift their fingers as rapidly as possible from the key whenever one of the designated (positive) lamps was turned on. They were warned to expect the occasional lighting of other lamps, but that reactions should not be made to these. The instructions also included a statement to the effect that if *S* found himself responding to the prohibited (negative) lamps, he was not to be too concerned but was to keep on reacting as quickly and accurately as possible.

After the instructions had been read aloud, *Ss* of both groups were given a series of 30 training trials, half of which involved the lighting of one positive lamp, and half the lighting of the

TABLE 1

PERCENTAGES OF RESPONSES MADE TO ALL SEVEN LAMPS BY *Ss* OF GROUPS 2–6 AND 3–5 IN EXP. 4 AND 5

(The 2–6 groups were instructed to react only to Lamps 2 or 6; the 3–5 groups were told to react only to Lamps 3 or 5.)

Exp.	Group	Lamp Number						
		1	2	3	4	5	6	7
4	2–6	45	100	36	40	46	100	52
4	3–5	38	47	100	71	100	52	37
5	2–6	58	100	53	39	52	100	55
5	3–5	28	25	100	66	100	34	27

other in a randomly determined sequence. Successive trials were given at intervals of from 10 to 15 sec., and the duration of the warning buzzer (2, 3, or 4 sec.), and hence also the duration of the foreperiod, was varied in a predetermined random order.

A test series, involving occasional presentations of negative lamps along with additional presentations of the two positive lamps, followed the initial 30-trial sequence without interruption or comment from *E*. The test series consisted of four presentations of each of the five negative lamps randomly interspersed among 53 additional randomized presentations of each of the two positive lamps. The number of consecutive positive-lamp retraining trials interpolated between any two successive negative-lamp test trials varied randomly between 2 and 10. With all *Ss* considered, no lamp appeared in any one testing position more frequently than in any other.

Two completely independent but identical experiments were run using this procedure, each with two groups of *Ss*. An interval of more than three years elapsed between the running of the two studies, and a different *E* served in each. To facilitate subsequent discussion, the three previous repetitions (1) of the single-light task will be designated hereafter as Exp. 1, 2, and 3. The two studies described here will be labeled Exp. 4 and 5.

Subjects.—In Exp. 4 and in 5, 25 volunteer students from the undergraduate psychology classes at Iowa served as *Ss* in each group. Thus for the two studies taken together the total number of *Ss* was 100, 50 per group. Each of the four basic groups contained approximately equal numbers of men and women.

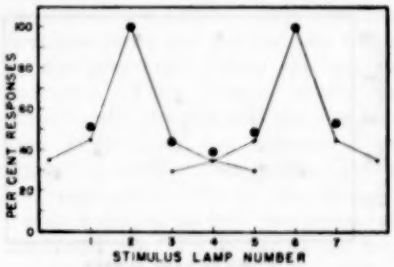

FIG. 1. The large solid dots show the percentage of responses made to all lamps during the test series by the 2–6 groups of Exp. 4 and 5 combined. Since *Ss* of these groups were told to respond to the lighting of either Lamp 2 or 6, response percentages are 100 at those points. The small dots and the connecting solid-line curves depict idealized, predicted gradients based upon empirical generalization data from studies in which *Ss* were instructed to respond to only one lamp.

RESULTS

Groups 2–6.—The instructions given these groups called for responses to Lamps 2 and 6 and prohibited responses to all other lamps. The percentages of responses to all lamps in Exp. 4 and 5 are shown in the first and third rows of Table 1. The values corresponding to the positive lamps, 2 and 6, are based on all presentations of those lamps to the 25 *Ss* of each group. Other entries are based on four presentations of each "prohibited" lamp to each *S*. The mean number of responses made to any of these lamps can be obtained by dividing the tabulated percentages by 25.

An examination of the data for the two independent experiments indicates reasonable comparability in view of the fact that the studies were run by two different *Es* and were separated by an interval of 3 yr. Statistical tests indicate that at only one point (Lamp 3) do the means for the two replications differ significantly

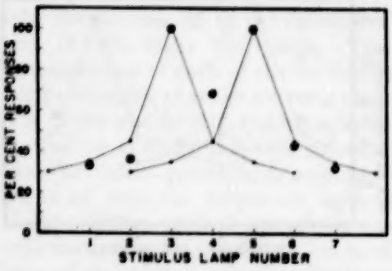

FIG. 2. The percentages of responses made by *Ss* of the combined 3–5 groups of Exp. 4 and 5 are indicated by the large solid dots. The solid-line curves are identical with those in Fig. 1. The summation effect is shown by the elevation of the point at Lamp 4 above the intersection of the solid-line gradients below it.

($P = .05$), hence the data from the two studies have been pooled. The resulting values are plotted in Fig. 1 as unconnected, solid circles.

The two double-winged, solid-line curves in Fig. 1 are *idealized* generalization gradients derived from the previously reported studies in which *Ss* were required to react to only the middle one of the seven lights. To obtain these idealized gradients, data from Exp. 1, 2, and 3 were first averaged to yield a single bidirectional gradient, then the right and left wings were averaged, and finally the resulting values were plotted in Fig. 1 on each side of Lamps 2 and 6. Such an average curve provides the best estimate of what the generalization curves should be like in the present experiments, on the assumption that the use of two lights rather than one and the use of lights other than the central one will not materially alter the gradients.

Examination of Fig. 1 reveals that these assumptions are reasonably well justified. When responses are required to either of two lights, the empirical points fall surprisingly close to the idealized single-light gradients. Lamp 4, which is spatially more remote from Lamps 2 and 6 than are Lamps 1, 3, 5, or 7, yields fewer responses than any of these latter four. It is also evident that Lamp 3, for example, which is three steps removed from 6, elicits approximately 15% more responses than would have been expected had Lamp 2 not been made positive. The same reasoning applies, of course, to the number of responses elicited by Lamp 5. Thus the data fit the pattern expected from the single-lamp studies, but since none of the unconnected points deviate significantly from the corresponding points on the solid-line curves, there is no evidence of summation.

Groups 3–5.—The response percentages obtained from the *Ss* of these groups, who were told to react only when Lamp 3 or 5 was turned on, are given in the second and fourth rows of Table 1. Here again not even the two largest differences between corresponding pairs of means are significant beyond the .05 level. The data have been combined, therefore, and plotted as the large solid dots of Fig. 2. The solid-line curves are identical to the idealized curves of Fig. 1.

The response percentages corresponding to Lamps 1, 2, 6, and 7 all fall reasonably close to the averaged single-light gradients, and even the most deviant one, that for Lamp 2, does not differ significantly from the idealized value. The figure for Lamp 4, however, has risen from an expected value of 45.0% (predicted from the single-light gradients) to 68.5%. This marked elevation of response frequency at the middle lamp constitutes the most significant finding of Exp. 4 and 5. It is a result that is clearly in accord with Hull's theorem concerning the summation of generalized reaction tendencies. The

fact that a significant summative effect was not obtained when the two positive lights were separated by 32° of visual angle (Groups 2–6), but was found when the angle was decreased to 16°, also coincides with Hull's prediction that the summative effect should be inversely related to the separation of the two "conditioned" points.

That the summation effect is probably not a chance phenomenon is indicated, in part, by the fact that the percentages of responses to Lamp 4 are very nearly the same for the two 3–5 groups in Exp. 4 and 5. Moreover, the tendency to react to Lamp 4 in both these groups is markedly facilitated in comparison to what it is in the corresponding 2–6 groups. To determine the statistical reliability of the summation effect a *t* test was run between the mean number of responses made to Lamp 4 (under instructions to respond to Lamps 3 and 5) and the mean number of responses made by all *S*s of Exp. 1, 2, and 3 to Lamps 3 and 5. This latter mean value provides the best available estimate of what the mean number of responses should be, in the absence of summation, to a negative lamp separated by one spatial unit (8°) from a positive lamp. In effect, this test involved a comparison of the central point of Fig. 2 with the intersection of the two idealized curves directly below that point. The results of this test proved to be highly significant ($t = 3.73$, $P < .01$).

Since the basic data from these experiments were accumulated during the relatively long series of test and retraining trials, the possibility exists that the summation effect increased over trials and hence was confounded by new learning. A rough check was made to test this possibility by making separate tabulations of the number of responses made to Lamp 4 on the first two times and the last two times it was presented during the testing sequence. These counts, made for Exp. 5 only, showed that the summation effect does not increase over trials but tends to disappear. Group 3–5 responded 72% of the time to Lamp 4 during its first two presentations and only 60% during its last two presentations. Group 2–6 yielded corresponding values of 30% and 28%. This finding accords well with the general observation characteristic of all these studies that responses to the prohibited lights are most likely to appear immediately after a long series of reactions to the nonprohibited lamps. Following the introduction of even a few test trials, however, and especially if some false responses occur thereon, the *S*s adopt a more cautious attitude which results in longer reaction times and fewer responses to the "generalized" stimuli.

Discussion

Although a phenomenon resembling the summation of generalized conditioned response tendencies has been observed in these studies, the authors are well aware that this does not constitute a direct test or verification of Hull's summation theorem. A direct test would require that the specific assumptions underlying the theorem be satisfied. For example, the positive reaction tendencies should be built up by the methods of classical conditioning rather than by verbal instructions. Moreover, Hull did not assume, in deriving his theorem, that extinction trials would be given to test stimuli lying between and beyond the conditioned ones. Our *S*s, however, were told not to react to the negative lamps, a procedure that is analogous, within the framework of classical conditioning, to the use of nonreinforced trials

to those stimuli. In spite of these genuine differences between our conditions and those demanded by Hull's theorem, the fact remains that there is a definite formal correspondence between our results and predictions from the theorem. This correspondence adds further weight to the suggestion that the extension of Hull's concepts to voluntary response situations may occasionally prove fruitful.

Whether tendencies to respond to the prohibited lamps in the present studies can be justifiably described as *generalized* tendencies is a question to which a universally convincing answer cannot be given. What can be said is that even when working with so-called voluntary responses, tendencies to respond to the prohibited stimuli do exist, and false (involuntary?) responses do occur. Moreover, the frequency of such reactions is partly dependent upon the spatial separation of the negative and positive stimuli. Regarded purely as empirical phenomena, the occurrence of false responses to negative stimuli and the occurrence of generalized conditioned responses are quite comparable. In studies of the generalization of classically conditioned responses, *S*s react to stimuli that have never been paired with the unconditioned stimulus; in the studies described here, not only do *S*s respond to stimuli to which training has not been given, but they do so even in violation of specific instructions to the contrary. Regardless of one's theoretical predilections, some mechanism must be invoked to account for these empirical findings. Hull's assumption that generalized habit strength mediates overt reactions to nonconditioned stimuli constitutes one such mechanism of reasonable utility.

Summary

Two identical experiments were conducted to determine whether tendencies to make generalized voluntary manual reactions combine in a manner predicted by one of Hull's theorems concerning the summation of generalized conditioned tendencies. In each experiment, *S*s were presented with a horizontally arranged row of seven stimulus lamps. Half of the *S*s were instructed to make a manual reaction to the lighting of either the second or the sixth lamp in the row. The other half were told to respond only to either the third or fifth lamp. Responses to all other lamps were prohibited by instructions. An initial period of training in responding to the designated lamps was followed by a test series in which occasional presentations of the prohibited lamps were interspersed among additional trials to the positive lamps.

Data obtained from the test series showed that *S*s instructed to react only to Lamps 3 or 5 made significantly more responses to the middle lamp (the fourth) than did *S*s required to respond only to 2 or 6. This increase in frequency of response to a negative stimulus lying between two positive ones, as a consequence of a reduction in the spatial separation of the positive stimuli, was interpreted as supporting Hull's concept of the summation of generalized excitatory tendencies.

References

1. Brown, J. S., Bilodeau, E. A., & Baron, M. R. Bidirectional gradients in the strength of a generalized voluntary response to stimuli on a visual-spatial dimension. *J. exp. Psychol.*, 1951, **41**, 52–61.
2. Gibson, E. J. Sensory generalization with voluntary reactions. *J. exp. Psychol.*, 1939, **24**, 237–253.
3. Hovland, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. genet. Psychol.*, 1937, **17**, 125–148.
4. Hull, C. L. The problem of stimulus equivalence in behavior theory. *Psychol. Rev.*, 1939, **46**, 9–30.

(Received February 26, 1955)

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

EXPERIMENTAL EXTINCTION AND DRIVE DURING EXTINCTION IN A DISCRIMINATION HABIT[1]

JOSEPH R. CAUTELA

Boston University[2]

Drive, as a variable determining performance, has received particular emphasis in recent years largely due to various theoretical speculations concerning the role of drive in behavior theory. A review of the studies concerning the relationship of drive to performance reveals the need for further systematic investigations of the effect of drive variation during training and during extinction.

In studies concerned with drive variation during extinction, extinction is carried out under various degrees of drive following acquisition under a single drive strength.

Using instrumental conditioning procedures (bar- and panel-pressing) Perin (**7**) and Horenstein (**4**) trained animals under 23-hr. food deprivation and extinguished different groups under various drive intensities up to 23 hr. Their studies show that the number of responses to extinction increases in a slightly positively accelerated manner as food deprivation intervals increased up to 23 hr.

In other studies, employing a slightly different design, Sackett (**8**), and Heathers and Arakelian (**3**) indicate that extinction rate is, to some extent, a function of the drive employed in training or in a previous extinction session.

The present study differs from previous studies in that: (*a*) a discrimination task is used, (*b*) deprivation intervals substantially above and below the deprivation interval used during the training procedure are employed, (*c*) deprivation intervals longer than 23 hr. are used.

METHOD

Subjects.—The *S*s were 70 experimentally naive, male albino rats approximately 90 days old at the beginning of the experiment. The animals were of the Sprague-Dawley strain and were obtained from the Charles River Breeding Laboratories.

Apparatus.—The discrimination apparatus is essentially the same as the one previously used by Spence (**10**). Briefly, it consists of a turntable divided into two platforms by means of a partition. The turntable rotated on a pivot made by two iron pipes. The smaller pipe, which was fastened to the turntable, fitted, with small tolerance, inside a larger one based on the floor.

The second main part of the apparatus consists of a stand which provided a one-way screen behind which *E* stood and which supported two elevated runways. One runway was painted a flat white and the other a flat black. Each runway was part of a unit which was connected to a food box. This food box consisted of a block of wood in which there was drilled a hole 1 in. in diameter. The combination runway and food cup fitted into a groove on which rested the base of the one-way screen. The weight of the screen further secured the runways in place. The runways were manipulated manually by lifting the one-way screen and pulling the runways out of the grooves (10 in. apart). After the manipulation of the runways, the turntable was rotated so the front (runway side) of the turntable which contained the *S* became the back turntable. In addition to the black and white runways, two duplicate runways painted an intermediate gray were used in preliminary training. The manipulation and use of the runways will be explained more fully in the preliminary training and the training sections.

Maintenance schedule.—On each of the first

[1] This article is based on a dissertation submitted to the Graduate School of Boston University in partial fulfillment of the requirements for the Ph.D. degree. The author wishes to express his appreciation to Dr. Leo J. Reyna for his advice and assistance during the course of this research.

[2] The author is now at Boston College.

seven days all *Ss* were given 12 gm. of Purina Lab Chow in individual feeding cages, and allowed to eat for 1 hr. This occurred at the same time each day. Water was available in the home and individual feeding cages at all times.

Habituation.—One hour before feeding time, on Day 8, *Ss* were placed on the front of the turntable, three at a time, with two gray runways in place. Sixteen food pellets (5.5 mm., 94 mg.) were scattered about the front turntable and five pellets were put in each cup. The *Ss* were allowed 30 min. to explore the apparatus and eat the food.

On Days 9 and 10 *Ss* were placed on the apparatus, one at a time, and allowed to eat five pellets from each gray food cup.

Preliminary training.—On Days 11 and 12 each *S* was given 10 trials on the apparatus with only one gray runway in place. After the *S* was placed on the back turntable the *E* went behind the one-way screen and raised the panel door which separated the back and front turntables. The *E* waited for the *S* to go up the runway, take a food pellet from the food cup, and go back to the front turntable. When *S* reached the front turntable, *E* rotated the turntable so that the front turntable, on which *S* was situated, became the back turntable. The *S* was now ready for the next trial. After the turntable was rotated, *E* manipulated the runway for the next trial. The runway was alternated in the following order: LRLRRLRLLR (Day 11); RLRRLRLLRL (Day 12).

Training.—On Day 13 the training for black-white discrimination began. The *Ss* were given 10 trials per day with both the black and white runways in place. Thirty *Ss* were always reinforced (found a pellet in the food cup) on the black runway and 30 *Ss* were always reinforced on the white runway. The noncorrection procedure was used under 23-hr. food deprivation and the reinforced alley shifted in the same order used in preliminary training. The criterion of mastery was 18 correct out of 20 trials, the last 10 being correct. All *Ss* were fed after reaching the criterion.

Extinction procedure.—After *S* reached the acquisition criterion, he was put into one of the following deprivation groups: 0, 6, 12, 47, and 71 hr. There were 10 *Ss* to a group.

The basis for inclusion in a particular deprivation group was the number of responses needed to reach the acquisition criterion. The first few *Ss* were randomly placed in different deprivation groups. The rest of the *Ss* were placed in particular deprivation groups to equalize the mean number of responses to reach the acquisition criterion for each deprivation group.

There was no food present in either food cup and the runways were alternated in the same order as in the training session.

The discrimination response was considered extinguished when the animal made no response for 3 min. A response choice was recorded when after raising the panel door, the *S* went on the runway and inserted his nose in the food cup.

Deprivation schedule for extinction.—The groups were fed according to the following plan: The *zero deprivation* group was fed 12 gm. for 1 hr. at 23 and 47 hr. after reaching the acquisition criterion. On the 70th hour *S* was put in the feeding cage for 55 min. with 12 gm. of food. At the end of the 55-min. period, if there was any food left in the feeding cage it was taken out and 12 additional gm. were put in the cage. If *S* had eaten all the food in the 55-min. period, an additional 12 gm. were put in the cage. The *S* was considered to meet the zero deprivation criterion when it did not eat any food for 3 min. following the introduction of the additional 12 gm. In no case did this take longer than 10 min. The *Ss* were then placed on the turntable and the extinction procedure was begun.

The feeding schedule of the other five groups is presented in Table 1. The extinction procedure was begun for all groups 71 hr. after reaching the acquisition criterion.

TABLE 1

NUMBER OF HOURS AFTER ACQUISITION THAT DEPRIVATION GROUPS WERE FED 12 GM. OF FOOD FOR 1 HR.

Group	Feeding at (Hrs.)			Test
0	23	47	70	71
6	23	47	64	71
12	23	47	58	71
23	23	47	—	71
47	23	—	—	71
71	—	—	—	71

RESULTS

Acquisition.—Bartlett's test (1) applied to the acquisition data shows homogeneity of variance. An analysis of variance applied to the data reveals no significant difference ($P < .05$) for the total number of responses or the correct number of responses to reach the acquisition

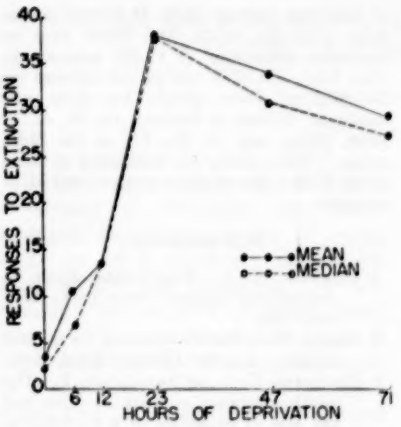

FIG. 1. The number of responses to extinction for the different deprivation groups.

criterion for the different deprivations groups. Likewise, the difference between the black-positive and the white-positive groups was not significant.

Extinction.—The curves for the mean and the median number of responses to extinction in Fig. 1 show an increasing function up to 23 hr. followed by a gradual decline to 71 hr. The curves up to 23 hr. are much steeper than the curves from 23 hr. to 71 hr., and the mean and median curves show very little difference in shape.

According to Bartlett's test (**1**) the assumption of homogeneity of variance was not met, even with the use of various transformations. As a result, the median test (**6**) was used to test the difference between the various groups. Chi-square results for differences between the groups are presented in Table 2. Fisher's exact method (**2**) and Yates's correction were applied to obtain the chi-square values. There were no significant differences ($P < .05$) between the 0-hr., 6-hr., and the 12-hr. groups nor between the 23-hr., 47-hr., and the 71-hr. groups. There were, however, significant differences, in every case, when the 0-hr., 6-hr., or the 12-hr. groups were compared with the 23-hr., 47-hr., or the 71-hr. groups.

DISCUSSION

The curves for the number of responses to extinction up to 23 hr. closely resemble those of Perin (**7**) and Horenstein (**4**), who obtained their results on instrumental type tasks.

In considering the curves beyond 23 hr. it would be reasonable to expect that animals which had been deprived of food for longer intervals would show greater persistence in making responses that one time led to food. But rather than a rise in the curves beyond 23 hr. we find that the curves gradually decline with 47-hr. and 71-hr. food deprivation. It is doubtful that this decline is due to the greater weakness in the organism resulting from longer food-deprivation intervals. Other studies (**9, 11**) indicate that this weakening effect does not occur until the fourth or fifth day of food deprivation.

TABLE 2

CHI-SQUARE RESULTS OBTAINED BY THE USE OF THE MEDIAN TEST FOR THE COMPARISON OF THE NUMBER OF RESPONSES TO EXTINCTION FOR THE VARIOUS DEPRIVATION GROUPS

Groups	χ^2	P
0 vs. 6	1.8	<.20
0 vs. 12	1.8	<.20
0 vs. 23	16.2	<.001
0 vs. 47	16.2	<.001
0 vs. 71	16.2	<.001
6 vs. 12	.2	<.70
6 vs. 23	9.8	<.01
6 vs. 47	11.8	<.001
6 vs. 71	5.0	<.05
12 vs. 23	7.2	<.01
12 vs. 47	5.0	<.05
12 vs. 71	5.0	<.05
23 vs. 47	.2	<.70
23 vs. 71	1.8	<.20
47 vs. 71	.2	< 70

The maximum resistance to extinction at 23 hr. is anticipated by Hullian theory. According to Hull (5) drive acts both as a multiplicative factor with habit strength to produce reaction potential, and as a cue stimulus. Thus if *S* is trained under a certain drive and subsequently performs under a stronger drive, this stronger drive, on the one hand, increases reaction potential (in its capacity as a multiplicative factor) but, at the same time, it operates to reduce the reaction potential since its cue value is a generalized one on the stimulus generalization gradient. This interpretation, of course, requires the assumption that the cue value of drive exceeds its energizing value. Support for this assumption is found in a study by Yamaguchi (13). He reports that when groups were trained at single deprivation intervals ranging from 3 hr. to 72 hr. and subsequently tested on all these points, response strength (as measured by reaction latency) was greatest when tested under the same drive used in the training procedure. Furthermore, from another report (12) based on the same study, Yamaguchi's results indicate that the function of the energy component of drive is **S** shaped.

Accordingly, the present total gradient can be accounted for by combining the following three assumptions: (*a*) the cue value of drive is greater than the energizing factor, (*b*) the gradient of the energizing factor is asymmetrical as proposed by Hull (5) in his formulation of stimulus intensity dynamism, (*c*) the gradient for drive as an energy component is **S** shaped on the basis of Yamaguchi's results.

Summary

The present investigation was concerned with the relationship between drive during extinction and rate of extinction of a discrimination task. Following the acquisition of a black-white discrimination response under 23-hr. food deprivation, six groups of 10 rats were extinguished under the following food-deprivation intervals: 0, 6, 12, 23, and 71 hr.

The mean and median number of responses to extinction increase up to 23 hr. and decline more gradually to 71 hr. There were no significant differences ($P < .05$) between the 0-hr., 6-hr., and 12-hr. groups, nor between the 23-, 47-, and 71-hr. groups, but there were significant differences between the 0-, 6-, or 12-hr. groups and the 23-, 47-, or the 71-hr. groups. These results are interpreted in terms of the joint action of drive as a cue and as an energizer.

References

1. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
2. Fisher, R. A. *Statistical methods for research workers.* London: Oliver & Boyd, 1948.
3. Heathers, G. L., & Arakelian, F. The relation between strength of drive and the rate of extinction of a bar-pressing reaction in the rat. *J. gen. Psychol.*, 1941, **24**, 243–258.
4. Horenstein, B. Performance of conditioned responses as a function of strength of their hunger drive. *J. comp. physiol. Psychol.*, 1951, **44**, 210–224.
5. Hull, C. L. Stimulus intensity dynamism (*V*) and stimulus generalization. *Psychol. Rev.*, 1949, **56**, 67–76.
6. Moses, L. E. Non-parametric statistics for psychological research. *Psychol. Bull.*, 1952, **49**, 122–143.
7. Perin, C. T. Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *J. exp. Psychol.*, 1942, **30**, 93–113.
8. Sackett, R. S. The effect of drive at the time of extinction upon resistance to extinction in rats. *J. comp. Psychol.*, 1939, **27**, 411–431.
9. Skinner, B. F. *The behavior of organisms.* New York: Appleton-Century-Crofts, 1938.
10. Spence, K. W. An experimental test of continuity and noncontinuity theories of discrimination learning. *J. exp. Psychol.*, 1945, **35**, 253–266.
11. Warden, C. J. *Animal motivation studies: the albino rat.* New York: Columbia Univ. Press, 1931.
12. Yamaguchi, H. G. Drive (*D*) as a function of hours of hunger (*h*). *J. exp. Psychol.*, 1951, **42**, 108–117.
13. Yamaguchi, H. G. Gradients of drive stimulus (S_D) intensity generalization. *J. exp. Psychol.*, 1952, **43**, 298–304.

(Received June 6, 1955)

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

DELAY OF REWARD AND PERFORMANCE OF AN INSTRUMENTAL RESPONSE [1]

GEORGE S. HARKER

State University of Iowa [2]

Within the framework of Hull's *Principles of Behavior*, (**2**) habit strength has the property of being built up as a relatively permanent state within the organism. In discussing the concept of habit Hull states: "... the immediate cause of habit-mediated action evocation must be a combination of (1) the stimulus event and (2) a relatively permanent condition or organization left by the reinforcement within the nervous system of the animal. This last is what is meant by the term habit" (**2**, p. 109). A parallel hypothesis presented in this formulation with regard to the acquisition of habits and the limiting values of habit strength that will be reached under particular experimental conditions is: "Habit strength at the limit of practice (m) will vary with the quality, as well as the quantity, of the reinforcing agent from a minimum of zero to a physiological maximum" (**2**, p. 129).

Two further statements concerning the effect of changes in the value of the reinforcing agent upon the maximum of habit strength are of concern to the experimentation to be presented. Although these statements refer specifically to magnitude or quality of reinforcement, in the context of their formulation, they may be considered to apply to delay of reinforcement also. "... according to the present hypothesis an increase in the amount of the reinforcing agent should be followed by a marked increase in the rate of habit strength acquisition; this, however, would gradually lessen as the new limit is approached, in accordance with the nature of the simple growth function" (**2**, p. 130). And, with respect to the converse condition, that of a reduction in the magnitude of reinforcement after training has progressed for a specified number of trials: "... ignoring secondary reinforcing effects, it is to be expected that successive reinforcements would result in a progressive weakening of the habit" (**2**, p. 131).

The apparent contradiction in the properties ascribed to habit strength is obvious. Habit strength cannot be an essentially permanent condition or organization that does not weaken or decrease at one moment and be called upon to decrease in amount with a decrease in the magnitude of reinforcement at the next.

More recently, Hull (**3**), in reformulating his principles, has changed his conception of the role of magnitude and delay of reinforcement in determining response strength. These two variables are now assumed to determine hypothetical incentive factors which act through a multiplicative relationship with habit to produce the excitatory potential of which the response is the directly observable consequent. This new formulation obviates the contradictions of the

[1] Abstracted from a thesis submitted to the Faculty of the Department of Psychology at the University of Iowa in partial fulfillment of the requirements for the degree of Doctor of Philosophy, 1950. The writer is deeply indebted to Professor Kenneth W. Spence.

[2] Now at the Army Medical Research Laboratory, Fort Knox, Ky.

earlier formulation and provides for the derivation of rather abrupt and extensive shifts in the level of response strength following either an increase or a decrease in the magnitude or delay of reinforcement.

The most relevant studies, the findings of which bear upon the hypothetical formulations as presented, are those of Crespi (1) and Zeaman (7). These studies provide data on acquisition and asymptotic performance values for animals trained under different magnitudes of food reward, as well as data on the effect upon level of performance of changing the magnitude of the food reward. The findings of these studies are in accord with Hull's newer formulation as to the magnitude of reinforcement in that the level of performance changed with a change in the magnitude of reinforcement.

The question may be asked: Does the effect of delay of reinforcement parallel that of magnitude of reinforcement? Experimental data is available on this question in the study by Perin (6) of bar-pressing in rats. Perin's findings support Hull's assumption that different limits of habit strength will be attained with different delays of reinforcement. Experimental findings on the effects of shifts in the delay of reinforcement are not in the literature except for the recent study of Logan (4) who utilized a relatively complex response incorporating approach, spatial discrimination, and bar-touching. Logan found both spatial preference and response speed to be responsive to changes in the delay of reward.

The present experiment was designed to study the course of acquisition of a simple instrumental response, bar-pressing, acquired under different delays of reward and to study the effect of shifting the delay of reward from a shorter to a longer delay and from a longer to a shorter delay upon the performance of the bar-pressing response.

Method

Subjects.—The *S*s were 141 naive hooded female rats from the colony of the Department of Psychology of the State University of Iowa. Their ages ranged from 90 to 100 days at the start of training.

Apparatus.—The apparatus was a modified Skinner box. The bar was rotated into the box by *E* through a slot in the wall of the box and rotated out by *S*'s response through a continuation of the slot in the floor. The feeding cup was immediately adjacent to the slot in the floor through which the bar rotated from the box. A ratchet prevented reverse rotation (lifting) of the bar. Provision was made for the delivery of a small lead shot to the under side of the food cup in combination with the food pellet. This device enhanced the auditory cue of the delivery of the reward and assured immediacy of reinforcement in spite of the ensuing search behavior to procure the food pellet. A 7-w. bulb located within the box served as a trial light. This light came on when the bar was rotated into the box by *E* and went off 30 sec. after the response was made. The sound of the ratchet on the bar provided a ready signal when *E* cocked the bar preparatory to rotating it into the box. The exclusion of irrelevant stimuli was accomplished by the use of a masking noise, a 60-cycle synchronous motor and gear train, rather than by the use of a semi-soundproof external shell.

The inside of the response box was 11 in. long, 7.75 in. wide, and 12 in. high. The bar entered the box 3.15 in. above the floor and was rotated by *E* to a position 2 in. above the floor. At this point the bar protruded into the box 1 in. The *S* was required to press the bar vertically downward .45 in. from which point a moderate spring action carried the bar down and out of the box. The force required to depress the bar was less than 1 gm.

The latency of the response was measured by a Standard Electric clock actuated by a microswitch applied to a cam on the bar assembly. The clock clutch was actuated when the bar was presented and released when *S* depressed the bar .15 in.

The delivery of the food pellet and lead shot at designated intervals following the response was accomplished by the use of a disc with appropriately spaced holes and a 2-rpm electric clock motor. The minimum delay possible,

immediate delivery, was approximately 1 sec. because of the starting lag of the motor and the fact that the disc could not be loaded at a point coincident with the delivery tubes.

Two additional boxes 8.5 in. long, 5 in. wide, and 10 in. high were used in the preliminary training. The first of these was a feeding box with wooden sides, screen top and floor, water bottle, and a food platform. The second had a solid floor, no water bottle, and was equipped with a feeding disc and food-delivery tubes similar to those of the experimental box. All boxes were duplicated in order to permit the running of two animals at the same time.

Preliminary training.—The *S*s were trained individually under 23-hr. food deprivation. A minimum of seven days was allowed for habituation to the experimental animal room and to establish the 23-hr. drive rhythm. Pretraining was accomplished in two phases during a 5-day period. Each *S* was allowed to eat (dry Purina Dog Chow pellets) in the first of the training boxes for 1 hr. on each of three days. Secondary reinforcement training was given in the box equipped to deliver food pellets in combination with lead shot on the remaining two days. The secondary reinforcement training consisted of 60 combinations of shot and pellet given in two groups of 30 each. The pellet size in the preliminary training was the same as that used in the experiment proper, approximately .05 gm.

Experimental procedure.—A group of *S*s was trained under each of five different conditions. For two control groups, Groups I and II, the time of delivery of the food pellet following the response was held constant at 1- and 10-sec. delay, respectively, throughout the course of the experiment. For the initial period of the experiment, Groups III and IV were trained under 1-sec. delay of food reward and Group V was trained with 10-sec. delay. After 40 trials, 30 on Day 1 and 10 on Day 2, the delays of reward for Groups III and V were interchanged. Group IV was shifted to no food reward or shot secondary reinforcement.

The data of the experiment were collected during 1-hr. sessions on each of the three days following the completion of the secondary reinforcement training. A dimly lighted experimental darkroom was used. Only indirect light entered the response boxes during the intertrial interval when the 7-w. trial light was off. The prototype trial sequence was as follows: The *E* loaded the food mechanism for the appropriate delay, cocked the bar, and rotated the bar into the box. If a response did not occur in 30 sec., it was elicited by holding a pellet of food in the slot above the bar. A 90-sec. intertrial interval was maintained except when it was necessary to elicit a response. Under these conditions, the 90-sec. intertrial interval was approximated as closely as *S*'s response would permit.

This procedure was continued throughout the three experimental days, each *S* being continued in the box to which it was assigned on the first day. Extinction trials differed from training trials in that the bar was rotated out of the box by *E* at the end of the 30-sec. period when no response occurred. In the extinction series, 50 presentations of the bar were made without food and shot, 20 on Day 2 and 30 on Day 3.

Thirty trials per day in the experimental training were taken as the practical working limit for massed trials without *S* having access to water. The pellet size was approximately .05 gm. which meant that *S* received 1.5 gm. of food in the experimental situation during a 45-min. period. The balance of the 7.5-gm. maintenance ration was given in the regular feeding cages immediately following the experimental session.

Subjects were assigned randomly to the initial training conditions of immediate reward (Groups I, III, and IV) and of delayed reward (Groups II and V). Within these conditions, assignment was random for Groups I, III, and IV while an effort was made to equate Groups II and V on the basis of the first day's performance. It was hoped by this means to stabilize the data of these groups with fewer *S*s. Twenty-one *S*s each were run in Group I, III, and IV; 39 *S*s each were run in Group II and V. The differential number of *S*s roughly equated the group standard errors.

Results

The main results of the investigation are presented to the same scale in Fig. 1 and 2. The point of change in the delays of reward is indicated by the vertical dotted line between Trials 40 and 41. In grouping the trials forward and backward from this point, the latency measure of Trial 41 was combined with those of Trials 40 and 39 as *S* had not yet experienced the new condition of delay. Thus, Trials 1, 2, and 90 became odd trials and were discarded. To avoid discarding the last trial of Day 1 and Day 2 and the first two trials of Days 2 and 3, the last trial of each day was combined with the first two trials of the next day and the curves plotted as continuous. Each plotted point repre-

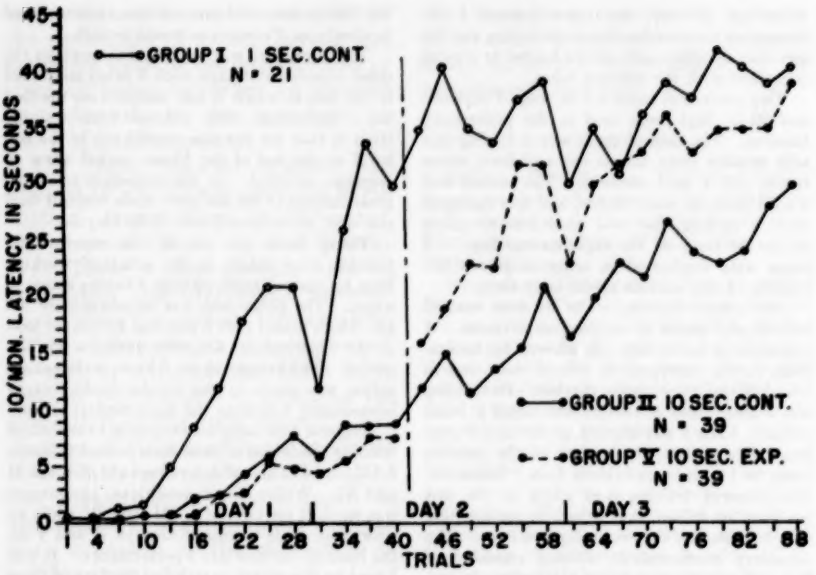

FIG. 1. Performance curves showing the effect of changing from 10-sec. to 1-sec. delay of reward.

sents the median for the group of the median response latency for each *S* on three successive trials. The performance measure plotted is the reciprocal of the response latency in seconds multiplied by 10. The original median latency measures for the several groups ranged from the arbitrary upper limit of 30 sec. to a minimum value of .24 sec. The shortest individual latency measured was .09 sec.

In Fig. 1, the solid lines are the learning curves for the controls, Group I and II. It is to be noted that the open circle curve for 1-sec. delay (Group I) rises more rapidly than the filled circle curve for 10-sec. delay (Group II). Furthermore, the curve for 1-sec. delay appears to have reached an asymptote by the end of Day 2, whereas, the curve for 10-sec. delay is still rising at the end of the training period.

In Fig. 1 and 2, the dotted lines give the effect on performance level of changes in the delay of reward. The result obtained for the shift from a less favorable reward condition (10-sec. delay) to a more favorable (1-sec. delay) is given in Fig. 1. This result is comparable to the results obtained with changes in magnitude and delay of reward in other experimental situations. Group V, the 10-sec. delay experimental group, showed a markedly greater gain following the shift to 1-sec. delay on Trial 41 than did Group II, the non-shifted 10-sec. delay control. By the end of Day 2 and throughout Day 3 of training, the performance of Group V was distinctly superior to that of its control, Group II, and approached in performance level that of the 1-sec. control, Group I.

The statistical significance of the mean performance gain on Day 2 by

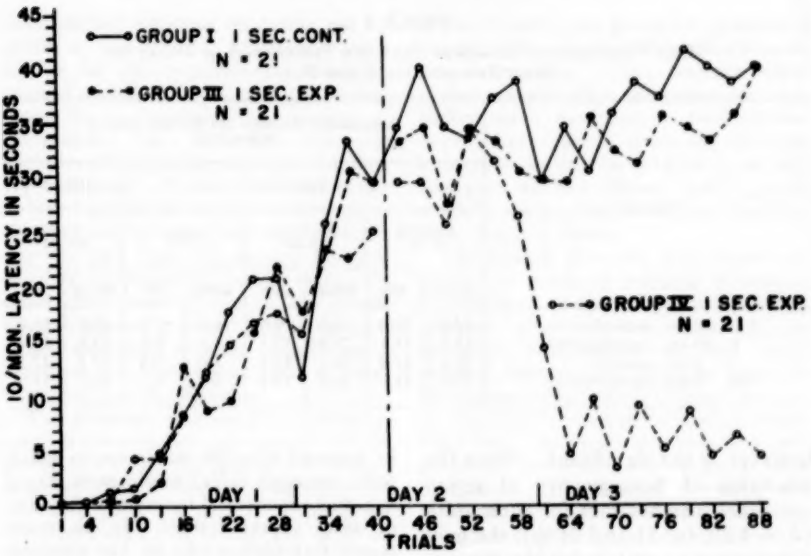

FIG. 2. Performance curves showing the effect of changing from 1-sec. to 10-sec. and from 1 sec. to infinite delay of reward.

Group V was ascertained by testing the hypothesis that this gain was not different from the mean gain evidenced by Group II. The relative gains were determined from Trials 31–41 to Trials 42–60, breaking the Day 2 data at the point of change in the delay of reward. The measure of group performance used in this test of the pre- and postchange periods was the mean of the *Ss*' reciprocal median response measure for the specified trials. These means are shown in Table 1. The mean gain difference between the gains of Groups II and V (8.4) is significant by the t test for unrelated measures ($t = 4.39$, $P < .01$).

The difference in level of performance of Groups II and V on Day 3 was also tested by a t for unrelated measures. The difference in performance level (9.4) was significant ($t = 2.77$, $P < .01$).

The result of the change in delay of reward from 1 sec. to 10 sec. is shown by the filled circle dotted lines of Fig. 2. Contrary to the findings for magnitude and delay of reward in other experimental situations, the shift to a less favorable delay of reward had little or no effect on the level of performance. By the end of Day 3, 49 trials after the shift, the curve for the shifted *Ss*, Group III, is intermingled with the curve for the control *Ss*, Group I. The mean performance measures for these groups pre- and postchange are also given in Table I. It will be seen from the table and also from Fig. 2 that Group III continued to improve on Day 2 following the change. This improvement continued through Day 3. The gain from pre- to postchange made by experimental Group III was not as large as the gain made by control Group I; the gain difference (1.6),

TABLE 1

MEAN PERFORMANCE MEASURES PRE- AND POSTCHANGE OF DELAY OF FOOD REWARD DAY 2 AND DAY 3

Group	Performance Measures (10/Median Latency in Seconds)							
	Day 2						Day 3	
	Pretrials 31–41		Posttrials 42–60		Gain		Trials 61–90	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
II. 10-sec. control	12.0	10.6	18.0	14.6	6.0	7.0	24.4	16.7
V. 10-sec. experimental	11.1	11.1	25.5	12.1	14.4	9.6	33.8	12.9
I. 1-sec. control	27.1	13.8	34.1	13.1	7.0	9.4	37.3	12.5
III. 1-sec. experimental	25.5	11.6	30.9	11.1	5.4	4.3	33.7	13.6

however, is not significant. Since the condition of homogeneity of group variances was not met for these data ($F = 4.89$, for 21 and 21 *df*), the gain difference was tested for significance by the Mann-Whitney (4) U test ($U = .78$, $P = .44$). The difference in level of performance (3.6) between Groups I and III on Day 3 was found not significant by a t for unrelated measures ($t = .88$, $P > .30$).

The effect of a change to infinite delay, no reward, following training with 1-sec. delay was demonstrated by the extinction of Group IV. Reward was withheld from this group following 40 training trials. The performance of Group IV is shown in Fig. 2 by the open-circle dotted line. It will be seen that despite discontinuance of the reward, performance continued to improve during the next 15 trials. Extinction of the bar-pressing response apparently began during the 19 unrewarded trials of Day 2; however, the major drop in performance level occurred on Day 3.

DISCUSSION

Hull has always assumed that delay of reward affects the asymptote of performance. In his earlier formulation (2) he assumed that the maximum to which habit strength (${}_sH_R$) would grow was a negative exponential function of the time of delay of the reward. In his more recent formulation (3), he has assumed that, not habit strength, but a new incentive factor, J, is a negative exponential function of the time of delay of reward. This factor is assumed to multiply habit strength to determine excitatory potential (${}_sE_R$). Thus, the new formulation provides for both shifts in level of performance with change in delay of reinforcement and differential terminal response levels for different delays of reward.

The present data (Fig. 1) suggest the possibility that for a simple instrumental response in an unrestricted situation, it is the rate of approach to the asymptote of performance which is affected by the delay-of-reward variable. The data of Perin (6), when replotted as reciprocals, support this observation in that the curve for his 10-sec. delay group rises more slowly than that for the zero delay group and rises continuously for the trials given. The possibility that both the asymptotic value and the rate of approach are affected is not precluded by either study for neither was carried sufficiently far to clearly ascertain the asymptote of the 10-sec. delay group.

The resistance to experimental extinction of Group IV suggests that the

instrumental response used was not as simple as was hoped. Apparently some feature of the experimental situation accrued sufficient secondary reinforcing properties to maintain improvement of the response for a short time following omission of the primary reward stimulus, the food pellet. Undoubtedly one of the factors responsible for the resistance of the response to experimental extinction was the fact that the amount of force required to depress the bar was extremely small (less than 1 gm.). The small amount of work involved in exerting such a force combined with the long (90-sec.) intertrial interval would lead to relatively little work inhibition.

The presence of secondary reinforcement in the situation, as revealed by the performance of the extinction group, suggests that secondary reinforcement probably played a part in maintaining the performance of Group III (Fig. 2) following the shift from 1- to 10-sec. delay of reward. Conceivably the immediacy of reward mediated by the secondary reinforcing agent (the *S*s frequently took the bar in their teeth or dug in the food cup) was sufficient to negate the shift in delay of reward. At the same time, the increased temporal delay of the primary reward stimulus would not alter the secondary reinforcing properties of the food-cup area in that it continued to be present, materially unchanged, at the time of arrival of the food.

The present findings with respect to the effect of delay of reward on performance are not in agreement with those of Crespi (**1**) and Zeaman (**7**) for magnitude of reward, and Logan (**4**) for delay of reward. These investigators found the limit of performance to be a function of the magnitude or delay of reward and that the limit was responsive to an increase or decrease in the magnitude or delay of the reward. One possible explanation for the absence of a decrement in performance of Group III following the shift to a less favorable reinforcement condition is that the 1-sec. group at the time of the shift was performing below the ultimate asymptote of the 10-sec. delay group and hence did not decrease. If the shift had been made at a later point in the training when performance of the 1-sec. group was at a higher level, the drop in performance might have been attained. Definite knowledge as to the asymptote of performance of the 10-sec. delay group would have provided a means of evaluating this hypothesis.

The Crespi, Zeaman, and Logan experiments involved learning situations in which a large proportion of the reaction sequence was approach, or approach and choice. In the unrestricted Skinner-box situation used in the present experiment the major portion of the reaction sequence was response initiation. Approach elements were progressively eliminated by *S*s as training progressed. The trained *S* during the intertrial interval characteristically assumed a position of expectancy by the food cup and the slot through which the bar was presented. The naive *S* or experimentally extinguished *S* explored the box and such irrelevant stimuli as it afforded. Consequently, much as the response measures reported portray an increase in the speed of the operationally defined response, bar-pressing in an unrestricted situation, they also portray a gross change in the physically apparent response or responses made by *S*. In this regard it is of interest to note that the rapidity of performance-level change in response to a shift from a more to a less favorable reinforcement condition in the studies noted (**1, 4, 7**) rank order consistently with the proportion of approach behavior in the response. That this is not true of the converse suggests the presence of a complicating factor, possibly secondary reinforcement, which is differentially affected by this dimension in these experimental situations.

Taken at face value the failure of a decrement in performance with the shift from the 1- to 10-sec. delay of reward is in agreement with Hull's earlier interpretation that habit strength is the factor affected by the delay-of-reward variable. Increasing the delay would not be expected to produce a decrement

in response since habit strength does not decrease or disappear. The increase in performance with the shift from a longer to a shorter delay is, also, in line with the earlier interpretation that the growth of habit strength is now to a new higher limit and hence the increments of growth will be greater than those of a group continued on the longer delay.

Summary

Five groups of rats were trained to depress a bar in an unrestricted Skinner-box situation under two conditions of delay of reward in order to study the effect of shifting the delay of reward once the training had progressed a specified number of trials. Three of these groups were subjected to shifts in the delay of reinforcement while two were continued unchanged to serve as controls.

Plots of reciprocal latency measures indicated that the rate of approach to the asymptote was affected by the delay-of-reward variable. The possibility that both the asymptote and rate of approach were affected, however, was not precluded.

A significant change in performance level was obtained for a shift in the delay of reward from a less favorable to a more favorable condition. The change in performance level was not significant for a shift in delay from a more favorable to a less favorable condition.

The response proved highly resistant to extinction, showing improvement for 15 trials following the change to nonreward. It is believed that this was due to the small amount of work involved in the execution of the response and to the presence of secondary reinforcement.

It is suggested that the deviate results of experimentation in this area might be equated through clarification of the role of secondary reinforcement in the various experimental situations. The results taken at face value suggest that Hull's original assumption that the maximum to which habit strength will grow is a function of delay of reward is correct, rather than his more recent formulation which specifies that a new incentive factor J is the one effected by time of delay.

References

1. Crespi, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, **55**, 467–517.
2. Hull, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
3. Hull, C. L. Behavior postulates and corollaries—1949. *Psychol. Rev.*, 1950, **57**, 173–180.
4. Logan, F. A. The role of delay of reinforcement in determining reaction potential. *J. exp. Psychol.*, 1952, **43**, 393–399.
5. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, 1947, **18**, 50–60.
6. Perin, C. T. A quantitative investigation of the delay of reinforcement gradient. *J. exp. Psychol.*, 1943, **32**, 37–52.
7. Zeaman, D. Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 1949, **39**, 466–483.

(Received May 23, 1955)

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

INFERENTIAL BEHAVIOR IN PRESCHOOL CHILDREN

HOWARD H. KENDLER AND TRACY S. KENDLER[1]

New York University *Barnard College*

This study was designed primarily to determine whether children, three to four years old, would be capable of inferential behavior. The experimental situation used was similar to the one suggested by Hull (1, p. 310) to serve as a test of his theoretical account of the reasoning behavior of rats as reported by Maier (3).

In our experimental procedure, which is diagrammatically represented in Fig. 1, *S* received three separate experiences. The *S* learned to pull A to get subgoal B, to pull X to get subgoal Y, and to pull B to get major goal G. In the test trial *S* was presented with A and X and told to choose the one that would obtain G. If he chose A, his behavior was designated as inferential.

In addition to discovering whether inferential behavior would occur, we were also interested in discovering whether the order of presentation of the three distinct experiences would influence the amount of inferential behavior. Hull's analysis, which presupposed the use of inarticulate *S*s, led to the prediction that order would influence the amount of inference. Reasoning that inferential behavior is mediated by the mechanism of the anticipatory goal response, Hull analyzed the process in the following manner: An r_g is developed to the major goal G which moves back to B and from there back to A. When presented with a choice between A and X, *S* should choose A because it evokes anticipatory goal responses for the major goal, which in turn produce cues which have become conditioned to responses in the direction of B and G. According to this analysis there should be more inference when the B-G experience precedes A-B than when it follows it, because in order for A to acquire the capacity to evoke an r_g for the major goal, B must have had an opportunity to become conditioned to the r_g of the major goal.

Although the present experiment was not primarily designed to test this hypothesis, and consequently does not provide a definitive test, it does throw some light on the effect of order of presentation of experiences on inferential behavior in preschool children.

METHOD

Experimental design.—The design included four experimental and four control groups with 16 *S*s in each group. Each experimental group had the three experiences in a different order. As shown in Table 1, B-G precedes A-B in two

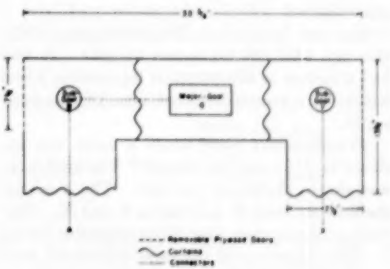

FIG. 1. Floor plan of apparatus.

[1] This experiment was supported by a grant from the National Science Foundation. The authors are indebted to Dr. M. R. D'Amato for his assistance in the design and construction of the apparatus. Acknowledgment is also made to Rolling Hills Nursery School, the Roslyn Nursery School and the Nursery School of the North Shore, all located in Nassau County, New York, for their kind cooperation in supplying *S*s and facilities for the execution of this research.

TABLE 1

DIAGRAMMATIC REPRESENTATION OF THE EXPERIMENTAL DESIGN

Groups	Order of Experiences			
	1	2	3	4
Experimental	A-B, X-Y, B-G	X-Y, A-B, B-G	B-G, A-B, X-Y	B-G, X-Y, A-B
Control	A-B, X-Y, G	X-Y, A-B, G	G, A-B, X-Y	G, X-Y, A-B

orders and follows it in two others. The control groups differed from the experimental groups only in that they attained G without ever experiencing B-G. These groups were included to control for possible effects, other than the B-G experience, that might influence *S*s choices during the test trials, e.g., the effect of which of the three experiences terminated the sequence.

Subjects.—The *S*s were 128 (68 M and 60 F) children between 34 and 59 mo. of age. They were drawn from three private nursery schools. Sixteen *S*s were assigned at random to each experimental group, with the limiting condition that eight *S*s in each group were drawn from children between 48 and 59 mo. old, and eight from children between 34 and 47 mo. old. The experimental groups were not equated for sex.

Apparatus.—The apparatus (Fig. 1) was a portable maze-like structure made of ¼-in. plywood in which the goals were pulled along the alleys by means of strings until they appeared at the appropriate openings. The entire apparatus was covered by a fitted plywood section so that *S* saw a gray box with two curtains at A and X. The covering plywood was jointed in a manner that permitted the top to be lifted to make the necessary adjustments of goal objects while screening *E*'s manipulation from *S*. At B and Y there were removable plywood "doors."

The minor goal objects at B and Y were either a small stuffed red ladybug or a small stuffed gray chicken. At G, the major goal, was a foreign toy sports car (Schuco-Akustico 2002) that ran when its brake was released. It also had a button in the center of its steering wheel which, when pressed, emitted a realistic car-horn sound.

Procedure.—In each school a room was assigned to *E* for the experiment. The apparatus was always placed on the floor. The *S* sat on the floor in front of and facing A and X. The *E* sat on the opposite side of the apparatus facing *S*. The experimental task was introduced to *S* as a game.

Each experimental *S* was presented with each experience twice in succession[2] but in different orders, as indicated in Table 1. For example, if *S* was in Exp. Group 1, *E* first put out the ribbon at A and told *S* to pull it. The *S* pulled and the stuffed ladybug emerged through the curtains at A. Although a string was clipped to both subgoals to prevent them from being pulled more than a few inches past the apparatus, there was sufficient play in it to allow *S* to handle the stuffed toy. After a few moments of handling, the ladybug was retracted by *E*. Then the ribbon was again put out past the curtains at A and the procedure described above was repeated. When the two A-B experiences were completed, the gold chain was set out past the curtains at X. The *S* pulled and out came the chicken, which was handled before it was retracted. This was repeated to provide two X-Y experiences.

For the B-G experience the ladybug was connected by a string to the car and the plywood door at B was removed. The *S* was instructed to come around to B and "peek in." He could see the ladybug but the car was hidden by the curtain between B and G. The *S* was then told to pull out the bug. When *S* pulled the bug, he also pulled out the car. He was permitted to play with the car for about one minute. During this time he made it go around the room once and tooted the horn two or three times. Then the ladybug and the car were replaced in the apparatus, hooked up again, and the procedure was repeated.

The next trial was the test trial. Both the ribbon at A and the chain at X were set out simultaneously. The *S* was instructed to pull *one* of them, the one that would get him the car.

Although the order of presentation of the experiences differed for the remainder of the experimental groups, each experience was conducted in precisely the manner described except that the side of the apparatus (right or left) on which A was placed, the subgoals (chicken or

[2] Preliminary research with this procedure suggested that one trial on each experience was not sufficient to obtain inferential behavior in 3- and 4-yr. children.

ladybug) and the character of the connectors at A and X (ribbon or chain) were counterbalanced for each experimental and control group.

The control groups differed from the experimental groups only in that the B-G experience was absent. In order to keep both the experience of pulling and the experience with the major goal constant, the control groups were presented with a small gray box, 8½ in. long × 3 in. high × 4 in. wide, in which the car was placed. Attached to the car was a green plastic string. The child pulled this and the car emerged from behind the curtains that covered the opening. He was permitted to play with the car in the same way that the experimental *Ss* were. The remainder of the procedure was identical with that of the experimental groups.

Results and Discussion

The test-trial results for all experimental and control groups are presented in Table 2. Evidently the experimental procedure is capable of producing inference in nursery school children. The difference in frequency between the total experimental and total control groups yielded a χ^2 of 8.29, $P = < .01$. There was no significant difference between the amount of inference demonstrated by the two age groups.

These results are in opposition to a conclusion reached by Maier based upon his study of reasoning in children which was reported in 1936 (**4**). His results showed that the preschool children in his experiment did not perform significantly better than chance when required to combine past experiences to reach a goal. Maier concluded that ". . . the ability to combine the essentials of two isolated experiences in such a manner as to reach a goal is rather late in maturing. It is rarely developed to a marked extent in children below six years of age." The experimental results on which Maier's conclusion was based are not necessarily inconsistent with the present results. The task Maier used was more complex than our relatively simple and direct procedure. The comparison between these two sets of data suggests the need to consider more carefully the nature of the experimental task before reaching such general conclusions.

Table 2 also shows that when either the A-B or X-Y experience immediately preceded the test trial, *Ss* tended to choose the opposite side on the test trial. This tendency occurred in both the experimental and control groups (3 and **4**).

Table 3 presents comparisons designed to explore the effects of the relative positions of B-G and A-B. In this analysis group results were combined (**2**) to provide a sufficiently large *N* for a valid test of statistical reliability. These results indicate

TABLE 2

Number and Percentage of *Ss* Who Made Inferential Choice on Test Trial for Each Experimental and Control Group

Group	Order	Experimental Groups		Control Groups	
		N	%	*N*	%
1	A-B, X-Y, B-G (or G)	12	75.0	7	43.8
2	X-Y, A-B, B-G (or G)	14	87.5	7	43.8
3	B-G (or G), A-B, X-Y	13	81.2	12	75.0
4	B-G (or G), X-Y, A-B	7	43.8	4	25.0
Total		46	71.9	30	46.9

TABLE 3

Number and Percentage of Inferential Responses on Test Trial According to Position of the B-G Experience

Position of B-G	Experimental Groups		Control Groups		χ^2	p
	N	%	*N*	%		
B-G (or G) last (groups 1 and 2 combined)	26	81.2	14	43.8	9.60	< .01
B-G (or G) first (groups 3 and 4 combined)	20	62.5	16	50.0	1.01	.50

that a significant amount of inference does occur when A-B precedes B-G but not when B-G occurs initially. These results fail to support Hull's hypothesis that a sequence of experiences in which B-G precedes A-B should produce more inferential behavior than when the order is reversed. We feel, however, that a more stringent test entailing direct comparisons between experimental groups, varying the B-G and A-B orders while *holding constant the last experience*, is required before a clear-cut conclusion can be drawn about the relevance of Hull's hypothesis to articulate organisms.

Summary

An experiment was conducted to determine whether children, 3–4 yr. of age, were capable of making an inferential response on the basis of past training. The experimental *Ss* received three separate experiences in a maze-like situation. Two of these experiences (A-B and X-Y) led to subgoals while the third experience (B-G) led to a major goal. The subgoal in the A-G experience served as the start of the B-G experience. The control *Ss* had a similar set of three experiences, with the exception that the experience involving G was not preceded by B. When *Ss* were given a choice between A and X with instructions to choose the alternative which led to G, significantly more experimental *Ss* than control *Ss* chose A. These results were interpreted as demonstrating inferential behavior in preschool children. The data also suggested that inferential behavior was more likely to occur when the A-B experience preceded B-G than when B-G occurred prior to A-B.

References

1. Hull, C. L. *A behavior system.* New Haven: Yale Univ. Press, 1952.
2. Lewis, D., & Burke, C. J. The use and misuse of the chi-square test. *Psychol. Bull.*, 1949, **46**, 433–489.
3. Maier, N. R. F. Reasoning in white rats. *Comp. psychol. Monogr.*, 1929, No. 29.
4. Maier, N. R. F. Reasoning in children. *J. comp. Psychol.*, 1936, **21**, 357–366.

(Received May 9, 1955)

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

TWO-CHOICE BEHAVIOR OF PARADISE FISH

ROBERT R. BUSH AND THURLOW R. WILSON

Harvard University[1]

Our problem stems principally from two experiments. Brunswik (**1**) observed the acquisition of a position discrimination by rats when food was placed more frequently in one box. Research by Humphreys (**9**) was comparable in that *S* had two choices with partial reinforcement of both. He required college students to guess on every trial whether or not a light would flash, and then in accordance with a predetermined schedule, the light did or did not flash. The Humphreys study exemplifies a noncontingent procedure for two-choice learning since the flash of the light did not depend upon the choice made by *S*. Brunswik's rats faced a contingent situation since the environmental change, presentation of food, was contingent in part on *S*'s response. A contingent two-choice research on humans has been performed by Goodnow (**2**, pp. 294–296). Her *Ss* decided on every trial which of two buttons to press. If the choice was correct, they earned a poker chip, otherwise not. Human two-choice learning with partial reinforcement has been further observed under contingent procedure (**3**) and under noncontingent procedure (**3, 4, 5, 6, 7, 8, 10**).

Bush and Mosteller (**2**) suggest that these two types of procedures are associated with different forms of asymptotic choice distribution (choice distribution after learning) for the individual *Ss*. In general, most *Ss* in a contingent experiment are found to have an asymptotic choice distribution of 100% selection of the favorable alternative. Noncontingent situations give rise to other kinds of choice distributions; in such experiments, the asymptotic proportion of choices of the favorable alternative has been observed to match the proportion of reinforcements scheduled for the alternative.[2]

We attempted to obtain the noncontingent results with nonhuman *Ss*. Red paradise fish were confronted by a position discrimination with partial reinforcement in which one side was correct a random 75% and the other side correct for the remaining 25%. The apparatus was a discrimination box with adjacent goal compartments. For the experimental *Ss*, *E* placed the food in the correct compartment regardless of whether *S* had entered the correct goal box; the division between the two goal boxes was transparent for the experimental group so that these *Ss* were able to see the food in the correct compartment when they had chosen incorrectly. The control group was run with an opaque divider separating the goal compartments in order to produce conditions comparable to those used by Brunswik.

Theory

We attempt to describe the experimental data within the framework of the

[1] This research was supported by the Laboratory of Social Relations, Harvard University. We are indebted to W. S. Verplanck for suggesting that we use fish in learning experiments and to F. Mosteller for numerous suggestions and criticisms.

[2] Besides contingent and noncontingent procedure, other kinds of factors, such as a gambling versus a problem-solving orientation, have been related to asymptotic choice distribution (**6**). We shall not deal with these other factors.

stochastic model given by Bush and Mosteller (2). On trial n (where $n = 0, 1, 2, \ldots$) there exists a probability p_n that S will choose the more favorable side. One of four events occurs on this trial and each leads to a different value of p_{n+1}. As in similar analyses, we assume that the effect of feeding is symmetrical for the two goal boxes; we make a similar assumption for nonfeeding. In addition, we assume that a long sequence of feedings on one side would tend to make the probability of going there unity. These special assumptions reduce the general model to the following statements about p_{n+1}.

Event	p_{n+1}	*Probability of occurrence*
favorable side, food	$\alpha_1 p_n + (1 - \alpha_1)$	$.75p_n$
favorable side, no food	$\alpha_2 p_n + (1 - \alpha_2)\lambda$	$.25p_n$
unfavorable side, food	$\alpha_1 p_n$	$.25(1 - p_n)$
unfavorable side, no food	$\alpha_2 p_n + (1 - \alpha_2)(1 - \lambda)$	$.75(1 - p_n)$

The model previously used for analyzing two-choice experiments using the contingent procedure is obtained from the above table by imposing the further restriction that $\alpha_2 = 1$. This assumption implies that nonfeeding is an event which does not alter the response probabilities. It was expected that this model would describe learning by the control group in the present experiment. Given this specific model, it can be shown that the asymptotic p for each S will be either 1.0 or 0; for the 75:25 schedule it is predicted that a high percentage of Ss will tend towards 1.0. The exact percentage depends upon the value of α_1.

We propose two specific models for the experimental group of the present experiment. These models are obtained from the foregoing table by imposing two different sets of additional restrictions which in turn are suggested by two different theories of learning. The first specific model, herein called the *information model*, is obtained by taking $\alpha_1 = \alpha_2$ and $\lambda = 0$. As a result, the first and fourth listed events in the foregoing table have the same effect on p_n; they correspond to food being placed on the favorable side. Similarly, the second and third listed events have the same effect; they correspond to food being placed on the unfavorable side. These restrictions appear to arise most readily from a cognitive learning point of view, because each trial may be described as providing information about the payoff schedules. This information model is equivalent to the models used by Bush and Mosteller (2) and by Estes (4) for describing human experiments with the non-contingent procedure.

The other specific model for the experimental group, herein called the *secondary reinforcement model*, is obtained from the additional restrictions, $\lambda = 1$ and $\alpha_2 > \alpha_1$. This model assumes that when S enters one goal box and sees food in the other goal box it is secondarily reinforced for the response just made. It has been shown that this model predicts that each S will have an asymptotic p of 1 or 0 and that more Ss will tend towards 1 than 0. The precise proportion that tend towards 1 depends on the values of α_1 and α_2.

We are chiefly concerned with predictions about the forms of the asymptotic distributions of choices of the favorable side. These predictions could be tested experimentally by running many trials in the experiment and obtaining a proportion of choices for each S during, say, the last 100 trials. The proportions thus obtained would form a distribution which could be compared with the predicted ones. Unfortunately, the mathematical analysis presented by Karlin (11) suggests that the convergence of the distributions of these models is very slow. Therefore, a great many trials would be required in the experiment to obtain the desired distribution. In view of these considerations, we are forced to examine the "near-

asymptotic" distributions. The information model predicts that such a distribution will be clustered around a point just below .75, whereas the secondary reinforcement model predicts that it will be U-shaped with a peak near 1 and a somewhat smaller peak near 0. The model for the control group ($\alpha_2 = 1$) also predicts a U-shaped near-asymptotic distribution, but the peak near 0 should be very small compared to that for the secondary reinforcement model. These predictions are compared with data below.

METHOD

Subjects.—The *S*s were 49 red paradise fish, 27 in the control group and 22 in the experimental group. The red paradise fish (*Macropodus opercularis*) is a hardy tropical fish about 2 in. in length selected because of its small demands for care. The *S*s were housed separately in tanks with a water temperature of 80° ± 1°F. This was the temperature indicated by our feeding studies for maximum appetite. Lighting was by fluorescent fixtures which were automatically turned on for a standard 12-hr. period each day to control the activity cycle. (This fish has a diurnal rhythm of activity.)

Apparatus.—The apparatus was a discrimination box as shown in Fig. 1. The maze was constructed of ¼-in. opaque white Plexiglas, except for parts of the goal boxes. The control group had a white opaque divider, whereas for the experimental group this divider was transparent. For one goal box the side opposite the entrance to the box was formed from a piece of opaque light yellow plastic; the corresponding side of the other box was white opaque. These sides could be interchanged. (Exploratory studies indicated that a position discrimination with identical goal boxes is learned very slowly by these fish.)

The apparatus was placed in a 10-gal. tank shielded from room lights. Lighting came largely from a 75-w. spotlight 2 ft. above the maze and focused on the start chamber. Care was taken to ensure that water conditions of this experimental tank were as close as possible to those of the home tanks of *S*s.

Feeding.—The experimental food was prepared fish eggs from an inexpensive (10 cents an ounce) caviar ("Lumpfish caviar" packed by Hansen Caviar Co., New York, N. Y.). These eggs were found to be a highly preferred food of the paradise fish and were convenient to obtain and store. The eggs were presented singly; the egg was held on the end of a medicine dropper

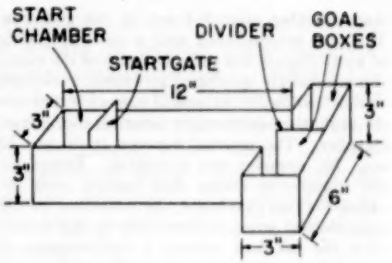

FIG. 1. Sketch of the discrimination apparatus.

by suction (the egg was 1 mm. to 2 mm. in diameter and larger than the opening of the dropper). To secure the egg, the fish was obliged to pull it from the dropper. A fish was required to earn all of its food by solving the discrimination problem.

Pretraining.—The pretraining took two or three days. On the first day the fish was fed eggs (10 or 20) by eye dropper in its home tank. For the next one or two days the fish underwent forced trials (10 or 20) in the maze. Half of the forced trials were to the right-side goal box. About one-third of the fish were rejected from the experiment at the end of pretraining or after one or two days of discrimination training leaving 49 *S*s. (Fish were rejected because they would not eat in the apparatus or because *E* made an error in procedure.)

Procedure.—All *S*s received a total of 140 trials, 20 trials a day or less. One goal box (the favorable side) was scheduled for reinforcement on 75% of the trials while the other goal was scheduled for reinforcement on the remaining trials. On a given trial only one goal box was correct. The trials for which the favorable side was incorrect were selected by restricted randomization within blocks of 20. The restriction was that runs of incorrect could not be longer than two. All fish had the same schedule.

The right, yellow side was favorable for about one-fourth of the *S*s; right, white for one-fourth; left and yellow for one-fourth, and left and white for one-fourth.

The procedure for the control group was as follows. The fish was released from the start chamber, and it swam down to the goal boxes. If the fish poked its nose into the goal box which was correct for that trial, *E* lowered a medicine dropper with a fish egg into the compartment (the dropper was secured to an arm) to allow the fish to feed. If the fish entered the incorrect goal box, no food was placed in the goal box. In either case, the fish was chased back into the

start chamber after 3–4 sec. in the goal box. This was accomplished with a piece of plastic of width slightly less than the width of the maze; the fish quickly developed avoidance tendencies to this "paddle." As soon as it was lowered into the tank, the fish promptly returned to the start chamber. The interval between trials was 12 sec. No retracing was permitted. Except for the transparent rather than opaque piece dividing the two goal boxes, the procedure for the experimental group differed only in one detail: after the fish had entered a compartment *E* placed the medicine dropper with a fish egg in the correct goal box. If the fish had entered the correct goal box, it secured the egg. Otherwise the fish could see the egg through the transparent divider but could not obtain it. Observations indicated that they did in fact see the egg on most of these trials.

Results and Discussion

Initial preferences.—Position and color preferences may strongly influence the results of a discrimination study. For this reason the balanced design described in the preceding section was used. This technique, however, tends to eliminate a *group* preference only. From an analysis of variance of the responses on the first 10 trials, we concluded that there were no group color or position preferences but that there were individual preferences. The stochastic models used in analyzing the data are sensitive to the entire distribution of initial probabilities, not only its mean. Therefore, it is necessary to consider the actual distribution.

TABLE 1

Observed Distribution of Choices of the Favorable Side During the First Ten Trials for the Two Groups of Fish Combined, and the Theoretical Distributions for the Binomial Model ($P = .5$) and for the Symmetric Beta Distribution with $s = .7$

Number Choices	Observed Number Fish	Predicted	
		Binomial	Beta
0	2	0.05	2.34
1	3	0.48	3.71
2	5	2.16	4.66
3	7	5.75	5.28
4	4	10.06	5.64
5	8	12.07	5.76
6	8	10.06	5.64
7	1	5.75	5.28
8	5	2.16	4.66
9	2	0.48	3.71
10	4	0.05	2.34
	49	49.07	49.02

One binomial observation for each initial probability is insufficient to determine anything about the initial distribution except the mean. Thus we must look at the number of successes (choices of the favorable side) by each fish during the first several trials and assume that the probability for each fish does not appreciably change during these trials. For this purpose the two groups of *S*s were combined, giving an *N* of 49, and the first 10 trials of the data were used. In Table 1 we show the frequencies of choice observed as well as those predicted by two models which are now briefly discussed.

The mean number of observed successes during the first 10 trials is .496 and so the balanced design accomplished its purpose. But, if we assume that each of the 49 fish had a binomial probability of .5, the predicted frequencies of choices are those shown in the third column of Table 1. The discrepancies are highly significant. The likelihood ratio test (**12,** p. 257) (this is essentially the chi-square test) leads to $P < .005$. Therefore we consider an alternative assumption: that the initial distribution is a symmetric beta distribution (**12,** p. 115) with a mean of .5. It may be written in the form

$$f(p) = C[p(1-p)]^s,$$

where C is a constant chosen so that the total density is unity, and where

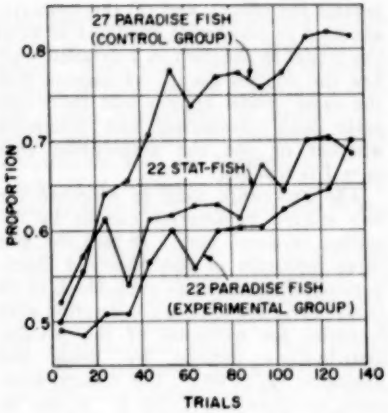

FIG. 2. Learning curve for each of the two groups of fish and for the 22 stat-fish which parallel the experimental group. Mean proportion of choices of the favorable side is plotted for each block of 10 trials.

s is a parameter which determines the spread of the distribution. The method of maximum likelihood (**12,** pp. 152–160) was used to estimate s from the data, giving .7 as the estimate. The distribution of successes during 10 trials can then be computed. The results are shown in the last column of Table 1 and the likelihood ratio test gives $P = .4$. This fit was considered satisfactory.

Learning curves.—In Fig. 2 we show the proportion of successes in blocks of 10 trials for each of the two groups of fish. It is clear that the control group learned more rapidly than the experimental group, but little more can be inferred from this figure. One can conjecture, of course, that the sight of food in the opposite goal box when food was not obtained slowed down the learning process. Just how this comes about can be determined only by a more detailed analysis of the data.

We hasten to note at this point that the models described above do not predict the relative rates of learning of groups of *Ss* run under different experimental conditions. Within the framework of the models, rates of learning are determined by the values of parameters which must be estimated from data. The models do predict, however, other properties of the data considered in the following sections.

The near-asymptotic distributions.—The two specific models for the experimental group—the information model and the secondary reinforcement model—make very different predictions about the shape of the distributions of successes after learning is nearly complete. In the second column of Table 2 we show the frequencies of successes during the last 49 trials (the number of successes varies from 0 through 49). The observed **U**-shaped near-asymptotic distribution is not determined by initial preferences alone; the rank-order correlation coefficient between the number of favorable choices on the first and last 10 trials is .22. The information model predicts a clustering

TABLE 2

DISTRIBUTION OF SUCCESSES (CHOICES OF THE FAVORABLE SIDE) DURING THE LAST 49 TRIALS FOR THE TWO GROUPS OF FISH AND FOR THE 22 STAT-FISH WHICH PARALLEL THE EXPERIMENTAL GROUP OF REAL FISH

Number Successes	Experimental Group	Stat-Fish	Control Group
0–4	4	4	1
5–9	1	2	0
10–14	2	0	0
15–19	0	0	1
20–24	0	0	2
25–29	0	1	0
30–34	1	0	2
35–39	2	2	3
40–44	2	3	7
45–49	10	10	11
	22	22	27

around 37 but this prediction is clearly *not* confirmed by the experimental group data. The secondary reinforcement model, on the other hand, predicts a U-shaped distribution with greater density at the high end than at the low end. This prediction *is* confirmed. On this basis alone we can choose the secondary reinforcement model in favor of the information model. Detailed questions of goodness of fit are considered in the following sections.

The model proposed for the control group involves the assumption that nonreward has no effect ($\alpha_2 = 1$) and it predicts that the near-asymptotic distribution of successes will also be U-shaped but with very small density at the low end. This indeed agrees with the data shown in the last column of Table 2; one out of 27 fish stabilized at the unfavorable side—it chose that side 46 times during the last 49 trials. The other 26 fish either stabilized on the favorable side or did not yet stabilize during the trials run. In the next section we consider the basic assumption that $\alpha_2 = 1$ made in the model for the control group.

Parameter estimates.—Having chosen the secondary reinforcement model for the experimental group, we need to estimate the primary reward parameter, α_1, and the secondary reward parameter, α_2. These estimates are required for two reasons: (*a*) we wish to measure the relative effects of primary and secondary reinforcement in this experiment (the smaller the value of α, the greater the effect), and (*b*) the estimates are used in measuring goodness of fit of the model to the data in a detailed way. For the control group, we assume that the same model applies and then estimate both parameters and determine whether or not the assumption that $\alpha_2 = 1$ is tenable.

TABLE 3

ESTIMATES OF THE TWO PARAMETERS OBTAINED FOR EACH OF THE TWO GROUPS OF FISH

Parameters	Experimental Group	Control Group
Primary reward, α_1	0.916	0.956
Secondary reward, α_2	0.942	0.986

The procedure used to estimate the two reward parameters cannot be described in detail here. It uses the first three moments of the observed distributions of successes in each block of 10 trials; these are used in conjunction with formulas for moments of the p-value distributions derived by Bush and Mosteller (**2**, p. 98). The results, however, are shown in Table 3. It can be noted that the secondary reward parameter, α_2, is larger for both groups than the corresponding primary reward parameter, α_1. This confirms the expectation that primary reward is more effective. (A small value of α implies a more effective event than does a large value.) For the control group, the value of α_2 is near 1.0 as assumed in the model for the control group, but the fact that it is not quite 1.0 suggests that nonreward is slightly reinforcing even for the control group. The result that α_1 is less for the experimental group than for the control group (primary reward more effective) is not predicted by any of the models.

The relative effects of primary and secondary reward for each group can be estimated as follows. We note that $(.916)^{.49} = .942$ and this means that secondary reward is about 60% as effective as primary reward for the experimental group. Similarly, $(.956)^{.31} = .986$, and so secondary reward is about 30% effective for the control group. These percentages may be in error appreciably because of the sampling errors in the parameter estimates, but they do indicate roughly the effects.

Stat-fish.—A convenient way of comparing model predictions with data is to run Monte Carlo computations or "stat-fish" as described elsewhere (**2**, pp. 129–131, 251–252). One hundred runs of 140 trials each were carried out on IBM

machines[3] using the parameter values given in Table 3 for the experimental group. From these 100 runs, a stratified sample of 22 runs was drawn such that the initial distribution of probabilities would approximate the symmetric beta distribution with the parameter $s = .7$. These 22 stat-fish can then be compared directly with the 22 paradise fish in the experimental group.

The "learning curve" of the stat-fish is shown in Fig. 2 along with those of the real fish. It can be seen that the stat-fish curve is slightly above the curve for the experimental group. This should not be interpreted as a discrepancy between the model and the data. Rather it is some indication of how well the model parameters were estimated from the data. Loosely speaking, the estimates were obtained by requiring that the learning rates of the model population and of the experimental sample be equal. To measure goodness of fit we must look at other properties of the data.

The near-asymptotic distribution of successes of the 22 stat-fish was obtained in the same manner as for the real fish. The results are shown in the third column of Table 2 and are sufficiently close to the corresponding frequencies of the experimental group that we consider formal tests for goodness of fit would be superfluous.

Many sequential properties of the data can be compared to the corresponding properties of the stat-fish "data" in order to obtain further measures of goodness of fit. Thus we have tabulated the distribution of runs (of successes and failures) for the experimental group and for the stat-fish. In Table 4 we show the mean and SD of the total number of runs, of the number of runs of various lengths, as well as the number of successes per S. It can be seen that all but one of the tabulated means are slightly smaller for the real fish than for the stat-fish, and that the variability of these measures is less for the real fish.

[3] We are indebted to B. P. Cohen and P. D. Seymour for making these computations.

TABLE 4

COMPARISON OF STATISTICS COMPUTED FROM THE DATA FOR THE EXPERIMENTAL GROUP OF 22 FISH AND FROM THE SEQUENCES OBTAINED FROM THE 22 STAT-FISH

Statistic	Experimental Group		Stat-Fish	
	Mean	*SD*	Mean	*SD*
Total number runs	27.3	13.7	29.8	20.2
Runs of length 1	12.9	6.7	14.2	10.1
Runs of length 2	4.4	3.8	5.3	6.3
Runs of length 3	2.0	1.8	3.0	2.6
Runs of length 4	1.5	1.6	2.2	2.2
Runs of length 5	1.1	1.3	1.0	1.4
Number successes	81.3	48.0	87.6	48.2

All these discrepancies are a result of the fact that two of the stat-fish never chose the unfavorable side and two others chose it only once each. These four stat-fish had initial success probabilities of .95, .95, .85, and .85, respectively. The smallest number of failures by the real fish is five. This suggests that better agreement would have been found if the initial distribution of probabilities had had less density in the extremes; the symmetric beta distribution was used only as an approximation to the true initial distribution. Furthermore, learning during the first 10 trials tends to spread out the distribution of response probabilities and so the true initial distribution probably had less variance than the symmetric beta distribution used in the stat-fish computations.

The distributions of the statistics given in Table 4 for the real fish and stat-fish can be compared in the same manner as used to compare two groups of Ss. The distributions are not normal and so we used the Mann-Whitney test (13). Comparison of each of the seven statistics listed in Table 4 led to P values greater than .3. Thus, we conclude that the model adequately describes much of the fine-grain character of the data.

Summary

A two-choice experiment designed to provide *Ss* with complete information about the outcomes of each choice on each trial is described. The *Ss* were 49 red paradise fish divided into two groups; the control *Ss* were run with the conventional procedure whereas the experimental *Ss* were given an opportunity to observe the presence or absence of food on both sides of the maze. Both groups were rewarded on one side 75% of the time and on the other side the remaining 25% of the time.

Two stochastic models for predicting the behavior of the experimental group are discussed. The "information model" assumed an increment in the probability of a fish choosing on a particular trial the side on which food was placed on the preceding trial. This model predicts that the distribution of choices approaches about .75 for all fish. The "secondary reinforcement model," on the other hand, assumes that sight of food in the opposite goal box reinforces the response just made and predicts that individual fish will approach 100% choice of one side or the other.

The data obtained support the secondary reinforcement model. Parameters which measure the effectiveness of primary and secondary reward are estimated from the data and then detailed comparisons between model predictions and experimental results are made. It is concluded that the secondary reinforcement model adequately describes much of the fine-grain structure of the data.

References

1. Brunswik, E. Probability as a determiner of rat behavior. *J. exp. Psychol.*, 1939, **25,** 175–197.
2. Bush, R. R., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.
3. Detambel, M. H. A test of a model for multiple-choice behavior. *J. exp. Psychol.*, 1955, **49,** 97–104.
4. Estes, W. K. Individual behavior in uncertain situations: an interpretation in terms of statistical association theory. In R. M. Thrall, C. H. Coombs, & R. L. Davis (Eds.), *Decision processes.* New York: Wiley, 1954.
5. Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47,** 225–234.
6. Goodnow, J. J., & Postman, L. Probability learning in a problem-solving situation. *J. exp. Psychol.*, 1955, **49,** 16–22.
7. Grant, D. A., Hake, H. W., & Hornseth, J. P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, **42,** 1–5.
8. Hake, H. W., & Hyman, R. Perception of the statistical structure of a random series of binary symbols. *J. exp. Psychol.*, 1953, **45,** 64–74.
9. Humphreys, L. G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, **25,** 294–301.
10. Jarvik, M. E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. exp. Psychol.*, 1951, **41,** 291–297.
11. Karlin, S. Some random walks arising in learning models I. *Pacific J. Math.*, 1953, **3,** 725–756.
12. Mood, A. M. *Introduction to the theory of statistics.* New York: McGraw-Hill, 1950.
13. Mosteller, F., & Bush, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology.* Cambridge, Mass.: Addison-Wesley, 1954.

(Received May 26, 1955)

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

STUDIES IN INCIDENTAL LEARNING: III. INTERSERIAL INTERFERENCE[1]

LEO POSTMAN AND PAULINE AUSTIN ADAMS

University of California

This experiment compares intentional and incidental learning with respect to interserial interference. We shall be concerned with the susceptibility of the two kinds of learning to interserial interference as well as with their relative effectiveness in producing interference.

Certain general predictions can be made on the basis of known differences between intentional and incidental learning. Let us consider susceptibility to retroactive inhibition (RI) first. Degree of original learning (OL) is typically higher under intentional than incidental conditions. Since degree of OL and amount of RI are inversely related (**4**), one might be led to predict less interference with intentionally learned material. We must, however, also take account of the distributions of habit strengths established under the two conditions of practice. The few items learned by incidental *Ss* are likely to have high associative strength. The items learned by intentional *Ss* are more numerous but include more items of low strength (**7**). Thus, the products of intentional learning may include a higher proportion of items susceptible to interference. We should, therefore, not expect major differences in the relative amounts of RI following intentional and incidental OL.

Consider next the effectiveness of interpolated learning (IL) under intentional and incidental conditions in producing RI. Amount of RI has been shown to vary directly with the number of interpolated items (**5, 10, 11**). Intentional *Ss* learn a larger number of items than incidental *Ss*. With OL held constant, intentional interpolation should, therefore, produce more RI than incidental interpolation.

Since amount of proactive transfer has in general been found to be a function of the same variables as amount of RI (**6,** pp. 394–399), we shall assume that the implications of our analysis for RI apply to proactive transfer as well. The present study investigates RI and associative inhibition (AI) under intentional and incidental conditions.

METHOD

Experimental design.—The experimental design (Table 1) comprises the conventional conditions for the determination of both RI and AI. There were 40 *Ss* in each of the groups. The *Ss* in the work groups learned the first list under either intentional or incidental instructions, then learned the second list under either intentional or incidental instructions, and finally were tested for free recall of either the first list (RI) or the second list (AI). Separate control groups were used for the RI and AI conditions. In the case of the RI groups, the rest interval was spent in reading selections from magazines.

Materials.—The two stimulus lists consisted of 20 nonsense syllables. In order to study the variable of familiarity of stimulus items, both lists comprised five syllables each of 0, 33, *66*, and 100% association value by Glaze's norms (**2**). One syllable of each association value appeared in each of five blocks of four syllables in the list. The syllables, printed in capital letters, were presented on 2 × 2 lantern slides. The height of the letters on the screen was 1¼ in.

Conditions of learning.—An orienting task was used to expose the stimulus lists under identical

[1] This research was facilitated by a grant from the Behavioral Sciences Division of the Ford Foundation.

TABLE 1

CONDITIONS OF PRACTICE USED IN MEASUREMENT OF RI AND AI
(A and B are stimulus lists. All tests for RI and AI were on List A.)

Groups	RI Conditions		AI Conditions	
	First List	Second List	First List	Second List
Controls	Intentional A	(Rest 4½ min.)	—	Intentional A
	Incidental A	(Rest 4½ min.)	—	Incidental A
Work groups	Intentional A	Intentional B	Intentional B	Intentional A
	Incidental A	Incidental B	Incidental B	Incidental A
	Intentional A	Incidental B	Intentional B	Incidental A
	Incidental A	Intentional B	Incidental B	Intentional A

conditions to intentional and incidental learners. Each syllable was exposed for 4 sec., after which *Ss* were given 6 sec. in which to write down a meaningful association suggested by the syllable. Intentional learners were, in addition, instructed to remember as many of the syllables as possible, although not necessarily in the order of presentation. In the work conditions, the second list was presented 30 sec. following the first. Instructions for the second series were given during the interval between lists. If the conditions for OL and IL were the same, the appropriate instructions were repeated. If there was a change from the incidental to the intentional condition, the instruction to memorize the syllables was added. If the change was from the intentional to the incidental condition, *Ss* were informed that the syllables in the second list were not to be memorized. Both the first and second list were presented once.

Retention tests.—A 5-min. test of free recall was given 30 sec. following practice on the second list. The *Ss* were encouraged to write down guesses as well as items of which they felt sure.

Subjects.—The *Ss* were 480 undergraduate students at the University of California. The *Ss* were run in small groups averaging six in number. Groups were assigned to conditions by rotation as they appeared in the laboratory, with a correction to equal *N*'s at the end. The *Ss* did not know the purpose of the experiment. A total of 50 *Ss* was discarded and replaced for the following reasons. Nine *Ss* in the RI groups and 10 *Ss* in the AI groups failed to follow instructions during the training trials. Four *Ss* in the RI groups and 5 *Ss* in the AI groups recorded their associations rather than the nonsense syllables on the recall test. Ten incidental *Ss* in the RI groups and 12 *Ss* in the AI groups stated that they had expected a test or had attempted to memorize the syllables during the presentation of the stimulus lists. About half of the *Ss* in this last category were from groups in which an incidental condition followed an intentional condition. Such an occasional generalization of the set was, of course, to be expected.

RESULTS

Degree of RI and AI.—Table 2 shows for all groups the average number of items correctly reproduced on the test of free recall and the percentages of RI and AI for the work conditions.

Comparison of the control groups shows that intentional *Ss* recalled more than incidental *Ss*. With an immediate test of recall (AI control groups), the intentional group significantly surpasses the incidental group ($t = 1.96$, $P = .05$). On the delayed test of recall (RI control groups), intentional learners again surpass incidental learners, but the difference is no longer significant ($t = 1.11$, $P = .28$). Previous studies (**1, 7**) have given some slight indication that forgetting may be more rapid following intentional than incidental practice. In the present study, the difference between the immediate and delayed control groups is significant for the intentional learners ($t = 2.77$, $P < .01$) but not for the incidental learners ($t = 1.37$, $P = .16$). To the extent that the

TABLE 2

MEAN RECALL SCORES AND AMOUNTS OF RI AND AI

Condition	Retroactive Inhibition			Associative Inhibition		
	Mean	*SD*	% RI	Mean	*SD*	% AI
Intentional Control	4.50	2.05	—	5.72	1.87	—
Incidental Control	3.98	2.14	—	4.72	2.61	—
Intentional-Intentional	1.65*	1.26	63.33	4.08*	1.85	28.67
Incidental-Incidental	2.18*	1.89	45.23	4.30	2.29	8.90
Intentional-Incidental	2.25*	1.62	50.00	3.88	1.79	17.80
Incidental-Intentional	1.30*	1.33	67.34	5.92	2.20	−3.38

* Significantly different from control group at the .01 level.

two types of practice are followed by differential rates of forgetting, the difference between the two groups of learners would diminish as a function of the time interval.

All RI work groups recall substantially less then their control groups. The differences between experimental and control groups are all significant at the .01 level. When the condition of interpolation is held constant, the products of intentional and incidental learning are equally susceptible to RI. There are, however, substantial differences in the amount of RI as a function of the condition of interpolation: for both types of OL, intentional IL produces more RI than does incidental IL.

In evaluating the statistical significance of the differences in the amount of RI, a procedure suggested by McAllister (**3,** p. 221) was used. The mean score of the appropriate control group is subtracted from each of the scores in the experimental groups, and the corrected scores of the experimental *S*s are used for purposes of tests of significance. An analysis of variance was performed on the corrected scores of the work groups. The results of the analysis are summarized in Table 3. The differences between the conditions of OL are not significant, whereas condition of interpolation is a highly significant source of variance. The interaction, OL × IL, is not significant, i.e., the effects of IL are parallel after intentional and incidental practice.

The percentages of AI are considerably smaller than the percentages of RI. When prior learning is incidental, AI is negligible or absent altogether. When prior learning is intentional, there are appreciable amounts of interference. The only work group which differs significantly from the control group, however, is the group which learns both lists under intentional instructions ($t = 3.99$, $P < .01$). An analysis of variance of the corrected scores of the work groups is summarized in Table 3. Again OL (second list) is not a significant source of variance whereas

TABLE 3

SUMMARY OF ANALYSES OF VARIANCE OF RETENTION LOSSES

Source	*df*	RI		AI	
		Mean Square	*F*	Mean Square	*F*
Original learning	1	3.78	1.54	.28	.07
Interpolated or prior learning	1	21.76	8.88*	51.75	13.62*
Interaction	1	.76	.31	5.27	1.39
Error	156	2.45		3.80	

* $P < .01$.

the condition of prior learning is highly significant. The absence of a significant interaction indicates that intentional and incidental *Ss* do not differ with respect to the detrimental effects of prior learning. However, the fact that the *Ss* who remain intentional learners throughout show the largest and only significant amount of AI, points to the possibility that the products of intentional learning may be more susceptible to AI than the products of incidental learning.

The retroactive and proactive effects are not directly comparable because of the difference in the time intervals between the end of learning and the test of retention. However, the differences between the two types of interference effects are so large that the results provide at least some indirect support for the conclusion that retroactive effects are greater than proactive effects for all combinations of intentional and incidental practice.

The effects of association value.— The test series included items differing widely in association value. Table 4 shows the percentages of RI and AI for items of high (100% and 66%) and low (0% and 33%) association value. The amount of RI for the two types of items varies as a function of OL and IL. With incidental OL, both kinds of interpolation produce a uniformly large amount of RI for items of low association value; RI for items of high association value is relatively small after incidental IL but rises markedly after intentional IL. With intentional OL, both types of items show a relatively moderate amount of RI after incidental IL. Interference for both types of items increases after intentional IL, but the rise is greater for items of low than high association value. Thus, for both groups of learners, items of low association value are more susceptible to interference than items of high association value. In agreement with earlier findings (**7**), we find the difference in favor of relatively familiar items more pronounced in the case of incidental learners. To determine the significance of this relationship, the difference between the corrected recall scores for the two types of items was determined for each *S*, and the difference scores were subjected to an analysis of variance. For both types of OL combined, the difference between retention losses for the two types of items is less after intentional than after incidental interpolation ($F = 8.16$, $df = 1$ and 156, $P < .01$). The interaction, OL × IL, is highly significant ($F = 10.97$, $df = 1$ and 156, $P < .01$). The effects of IL on the two types of items follow a different pattern for intentional and incidental learners.

TABLE 4

PERCENTAGES OF RI AND AI FOR ITEMS OF HIGH AND LOW ASSOCIATION VALUE

Condition	RI		AI	
	High	Low	High	Low
Intentional-Intentional	58.20	74.14	23.46	41.79
Incidental-Incidental	37.29	68.29	12.98	0
Intentional-Incidental	49.18	51.72	16.03	22.41
Incidental-Intentional	67.80	68.29	−2.47	−5.97

The picture is rather different for the AI groups. Only the group learning both lists under intentional conditions showed a significant amount of AI. For this group, AI is less for items of high than for items of low association value. The difference is in the same direction when intentional prior learning is followed by incidental learning. With incidental prior learning, the amounts of AI were negligible, and there are no systematic variations as a function of association value. Analysis of

variance fails to show any significant variation between conditions in the relative amounts of retention loss for items of high and low association value. Since there was reliable AI in only one case, significant effects were probably not to be expected.

Intrusions.—The average number of intrusions at recall for each of the work conditions is presented in Table 5. In general, the number of intrusions is small, and the variation among conditions is not extensive. The small number of intrusions is probably a function of the low degree of IL (1 trial). In the case of the RI groups, there appear to be more intrusions when both OL and IL are carried on under the same instructions than if the instructions are different. This last finding suggests that a common set at the time of OL and IL favors overt competition at recall (cf. **8**). Analysis of variance, however, fails to reveal significant variations in the number of intrusions as a function of the experimental conditions. Whatever significance we attach to variations in intrusions must be based solely on their consistency with other findings. There appears to be little systematic variation in the number of intrusions given by the AI groups, and again there are no statistically significant effects.

TABLE 5

MEAN NUMBERS OF INTRUSIONS AT RECALL

Condition	RI	AI
Intentional-Intentional	.88	.62
Incidental-Incidental	.82	.70
Intentional-Incidental	.75	.75
Incidental-Intentional	.56	.60

DISCUSSION

The results on interserial interference are in agreement with our expectations based on the difference between the distributions of habit strengths established under intentional and incidental conditions. We have assumed that the proportion of associations sufficiently strong to resist interference will be higher for incidental than for intentional *S*s. On the other hand, intentional learning produces stronger links among the individual members of the series. The net result is that intentional and incidental learning are found to be equally susceptible to RI. Intentional IL, however, produces more RI than does incidental IL. This result was expected on the basis of the known relationship between the sheer number of items learned during IL and the amount of RI. Intentional interpolation may, in addition, induce a greater readiness to use items from the second list as overt responses. Our findings on intrusions give only marginal support to this interpretation. These conclusions must for the present remain limited to low degrees of OL and IL such as were used in this study.

It should be noted that in a study reported by Prentice (**9**) incidental learners were found to be more susceptible to RI than intentional learners. However, Prentice based his conclusion on a comparison of the two work conditions without taking account of the higher scores of the intentional control group. When the scores of the experimental groups are corrected by subtraction of the mean control scores, according to the procedure described above, there is no significant difference between the two groups of learners ($t = 1.14$). There is, therefore, no essential disagreement between our findings and those of Prentice. Since Prentice pooled the results for *S*s who had intentional and incidental IL, the results of his study and ours cannot be compared with respect to the effects of different conditions of interpolation.

The results of the AI conditions in general support the conclusions drawn from the data on RI. Intentional learning is again a more serious source of interference than incidental learning. There is also some indication that

intentional learners are more susceptible to AI than incidental learners. In view of the fact that AI is generally low and variable, however, this difference should not be unduly stressed. The important fact from the point of view of the present analysis is that intentional learners again fail to show greater resistance to interference than incidental learners.

Summary

This experiment is concerned with RI and AI under intentional and incidental conditions. The materials were lists of 20 nonsense syllables. The work conditions included all possible combinations of intentional and incidental practice of the two lists. All RI work groups showed significant amounts of interference. Intentional IL produced more RI than incidental IL, but relative amount of RI was independent of the condition of OL. AI was significant only if both lists were learned intentionally. Intentional practice resulted in the learning of a larger number of items during interpolation and hence was a more effective source of interference. Materials learned selectively under incidental conditions were at least as resistant to interference as intentionally learned materials.

References

1. Biel, W. C., & Force, R. Retention of nonsense syllables in intentional and incidental learning. *J. exp. Psychol.*, 1943, **32,** 52–63.
2. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35,** 255–267.
3. McAllister, D. E. Retroactive facilitation and interference as a function of level of learning. *Amer. J. Psychol.*, 1952, **65,** 218–232.
4. McGeoch, J. A. The influence of degree of learning upon retroactive inhibition. *Amer. J. Psychol.*, 1929, **41,** 252–262.
5. McGeoch, J. A. Studies in retroactive inhibition: VII. Retroactive inhibition as a function of the length and frequency of presentation of the interpolated lists. *J. exp. Psychol.*, 1936, **19,** 674–693.
6. McGeoch, J. A., & Irion, A. L. *The psychology of human learning.* New York: Longmans, Green, 1952.
7. Postman, L., Adams, P. A., & Phillips, L. W. Studies in incidental learning: II. The effect of association value and of the method of testing. *J. exp. Psychol.*, 1955, **49,** 1–10.
8. Postman, L., & Postman, D. L. Change in set as a determinant of retroactive inhibition. *Amer. J. Psychol.*, 1948, **61,** 236–247.
9. Prentice, W. C. H. Retroactive inhibition and the motivation of learning. *Amer. J. Psychol.*, 1943, **56,** 283–292.
10. Twining, P. E. The relative importance of intervening activity and lapse of time in the production of forgetting. *J. exp. Psychol.*, 1940, **26,** 483–501.
11. Underwood, B. J. The effect of successive interpolation on retroactive and proactive inhibition. *Psychol. Monogr.*, 1945, **59,** No. 3 (Whole No. 273).

(Received March 31, 1955)[1]

[1] Editor's note.—Publication of this article was delayed for two months in order to print it in the same issue with IV and V in the series, which follow it in this issue.

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

STUDIES IN INCIDENTAL LEARNING: IV. THE INTERACTION OF ORIENTING TASKS AND STIMULUS MATERIALS[1]

LEO POSTMAN AND PAULINE AUSTIN ADAMS

University of California

In experiments on incidental learning an orienting task must be used which ensures *S*'s exposure to the stimulus materials in the absence of instructions to learn. The intentional learners, with whom the incidental learners are compared, may or may not be given this orienting task. The degree of difference between the two groups depends on whether or not the intentional learners perform the orienting task assigned to the incidental learners. As Saltzman (**6**) has shown, the usual difference in favor of the intentional learners may be drastically reduced when both groups perform the orienting task. The performance of the orienting task may interfere with those responses to the learning materials which are essential for the formation of associations. If only the incidental learners perform an orienting task, the superior performance of the intentional learners may be ascribed to *both* the intention to learn *and* to the freedom from activities interfering with the formation of associations. It is necessary, therefore, to take account of the effects of the orienting task in evaluating the difference between intentional and incidental learners.

The activities required by different orienting tasks may be more or less compatible with the formation of associative connections. Let us assume that both intentional and incidental learning depend on the performance of differential responses to the stimuli during exposure to the learning materials (**4, 5**). The activities required by the orienting task may facilitate or inhibit the performance of these differential responses to varying degrees. The magnitude of such effects should also vary with the nature of the stimulus materials. Meaningful items evoke differential responses more readily than do nonsense items. Differential responses made to meaningful items should, therefore, be less susceptible to interference from the orienting task. Hence we would expect the learning of meaningful materials to vary less as a function of the orienting task than the learning of nonsense materials. By the same token, the more meaningful the materials the less the difference between intentional and incidental learners should be influenced by the nature of the orienting task.

In the present experiment, orienting tasks were used which favored the differentiation of the learning items to varying degrees. The effects of these tasks on intentional and incidental learners were investigated with meaningful and nonsense materials.

METHOD

Materials.—Two sets of stimulus materials were used: a list of nonsense syllables and a list of adjectives. The nonsense list consisted of 30 syllables with association values of 40 and 46.67%. (**1**). The list of meaningful items consisted of 30 adjectives taken from the list published by Haagen (**2**). For both lists, there

[1] This research was facilitated by a grant from the Behavioral Sciences Division of the Ford Foundation.

were two different random orders used equally often under each experimental condition.

Orienting tasks.—There were two orienting tasks which were designed to interfere with learning to different degrees: (*a*) giving meaningful associations to the items, (*b*) matching the items with geometric figures.

The Association Groups were instructed to give a meaningful association to each of the items. Each item, printed in capital letters, was presented on a separate page of a booklet. The *Ss* turned the pages of the booklet at the rate of one every 10 sec. and wrote their associations in a space next to the item. The Figure-Matching Groups matched each item to a geometric design. The *Ss* were instructed to select one of the following geometric designs as best "fitting" a given item: square, triangle, circle, cross, arrow, diamond, and wavy line (sine curve). Drawings of these figures were in full sight of *Ss* throughout the experiment. The instructions suggested that better than chance agreement is often found for this type of intuitive judgment. The instructions also emphasized that the judgments should be made as quickly as possible and without regard to the meaning of the item. The method of presentation was the same as that used with the association task.

The association task was expected to be more compatible with learning than the figure-matching task. The task of giving meaningful associations to the items required the type of differential response which is assumed to be conducive to the formation of associations. However, the performance of the orienting task may interfere with the formation of intraserial linkages, e.g., by preventing the rehearsal of sequences of items. The task of matching geometric figures with the items did not encourage the type of differential responses which are favorable to associative learning. On the contrary, the instructions emphasized the need for quick and immediate reactions to the stimuli. Moreover, the pairing of different items with the same geometric designs may have resulted in response-produced generalizations among the members of the series and thus have retarded their differentiation.

Experimental design.—A group of intentional learners and a group of incidental learners were used for each combination of materials and orienting tasks. With two types of materials, two orienting tasks, and two conditions of practice, the total design comprised eight experimental groups. There were 32 *Ss* in each of the groups. After being introduced to the orienting task, the intentional learners were given the additional instruction that their memory for the items would be tested at a later time. They were requested to learn the items without regard to the order of presentation. In all other respects, intentional and incidental learners were treated alike.

The experimental design also included two Intentional Control Groups. One of these control groups learned the list of nonsense syllables; the other group learned the list of adjectives. The control *Ss* did not perform an orienting task but were merely instructed to learn the items without regard to the order of presentation. The manner and rate of stimulus presentation were the same as for the groups performing an orienting task.

Recall test.—A 5-min. test of free recall was given immediately after presentation of the stimulus series. The *Ss* were requested to write down as many items as they could remember, without regard to the original order of presentation. They were encouraged to put down their guesses if they were not sure about an item.

Subjects.—The *Ss* were 320 undergraduate students at the University of California. The groups learning nonsense syllables and the groups learning adjectives were tested some months apart.[3] In each half of the experiment, the *Ss* were assigned to the conditions in rotation as they appeared in the laboratory, with correction for equal *N*'s. The *Ss* served in small groups, averaging four in number. None of the *Ss* learning adjectives had to be discarded. Seven *Ss* learning nonsense syllables had to be discarded and replaced. Five *Ss* (four intentional and one incidental) failed to follow instructions during the presentation of the materials. One incidental *S* recorded his associations rather than the syllables on the test of recall, and one incidental *S* anticipated the test of recall.

Results and Discussion

Table 1 shows the average number of nonsense syllables recalled by the various groups. Consider the intentional groups first. The Intentional Control Group surpasses both intentional groups performing an orienting task. Analysis of variance shows the differences among the intentional groups to be highly significant ($F = 12.82$, $df = 2$ and 93, $P < .01$). According to Tukey's gap test (7),

[3] This time interval is due to a revision in procedure which made it desirable to replace the groups learning nonsense syllables.

TABLE 1

AVERAGE NUMBER OF ITEMS RECALLED BY THE VARIOUS GROUPS

Orienting Task	Nonsense Syllables				Adjectives			
	Intentional		Incidental		Intentional		Incidental	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Control	7.22	2.71	—	—	13.62	5.58	—	—
Associations	5.66	2.31	3.59	2.06	13.94	2.95	11.56	2.60
Figure matching	3.09	2.11	3.03	2.58	12.72	2.31	10.03	2.69

the breaks between successive means are significant at the .01 level. Both orienting tasks reduce the amount of intentional learning, but the figure-matching task is significantly more detrimental than the association task. With nonsense materials, the effectiveness of intentional learning depends on whether *Ss* perform an orienting task, and if they do, on the specific nature of the task. The difference between intentional and incidental learners is, in turn, a function of the orienting task.

Comparing intentional and incidental learners, we find (*a*) a relatively large difference between the two Association Groups, and (*b*) virtually no difference between the two Figure-Matching Groups. To test the significance of these relationships, the recall scores of the experimental groups were subjected to a two-way analysis of variance which is summarized in Table 2. The over-all difference between intentional and incidental learners is significant, as is the difference between the two orienting tasks. In addition, there is a significant interaction between instructions and orienting tasks, i.e., the difference between the two groups of learners depends on the nature of the orienting task. The more serious the interference from the orienting task, the smaller the difference between the two groups of learners becomes.

The mean recall scores for adjectives are also presented in Table 1. Under all conditions, recall is considerably higher for adjectives than for nonsense syllables. The differences among the three intentional groups do not reach significance ($F = 1.34$), although the Association Group again shows higher recall than the Figure-Matching Group. With meaningful materials, the performance of an orienting task has only limited effects on intentional learning. The nature of the orienting task does not influence the difference between intentional and incidental learners. An analysis of variance of the recall scores of the experimental groups is summarized in Table 2. As in the case of nonsense syllables, there is a significant over-all difference between intentional and incidental learners. For both groups combined, the association task results in significantly higher retention than the figure-matching task. The interaction between instructions and orienting tasks is, however, clearly not significant. In contrast with the results obtained for nonsense syllables, the intentional *Ss* retain their superiority over incidental *Ss* even under conditions unfavorable to learning.

Our results lead to the general con-

TABLE 2

SUMMARY OF ANALYSES OF VARIANCE OF RECALL SCORES

Source	*df*	Nonsense Syllables		Adjectives	
		Mean Square	*F*	Mean Square	*F*
Instructions	1	36.13	6.77*	205.03	28.09**
Tasks	1	78.13	14.63**	60.50	8.29**
Instructions × Tasks	1	32.00	5.99*	.78	.11
Error	124	5.34		7.30	

* $.01 < P < .05$.
** $P < .01$.

TABLE 3

SUMMARY OF OVER-ALL ANALYSIS OF VARIANCE

Source	*df*	Mean Square	*F*
Instructions	1	25.21	23.34**
Tasks	1	24.54	22.72**
Materials	1	613.95	568.47**
Instructions × Tasks	1	2.44	2.26
Instructions × Materials	1	.65	.60
Tasks × Materials	1	3.31	3.06
Instructions × Tasks × Materials	1	4.87	4.51*
Error	248	1.08	

* $.01 < P < .05$.
** $P < .01$.

clusion that the difference between intentional and incidental learners is a joint function of the nature of the learning materials and the orienting task performed during exposure to the stimulus items. With relatively difficult nonsense materials, intentional *Ss* surpass incidental *Ss* only as long as there is no serious interference from the orienting task. With the easier meaningful materials, the differences between the two groups of learners remain independent of the nature of the orienting task. The significance of the relationships between instructions, orienting tasks, and stimulus materials was tested by subjecting the results for both types of materials to an over-all analysis of variance (Table 3). The results of all groups except for the two Intentional Control Groups were included. Since variances as well as means were higher for adjectives than for syllables under all conditions, the square-root transformation recommended by Mosteller and Bush (**3**, pp. 326–328) was used. The main effects—instructions, tasks, and materials—are highly significant. The only other significant source of variance is the higher-order interaction, Instructions × Tasks × Materials. The significance of this interaction directly confirms the conclusion that the difference between intentional and incidental learners depends on both the meaningfulness (difficulty) of the materials and the nature of the orienting task.

SUMMARY

In experiments on incidental learning an orienting task is used to expose *Ss* to the stimulus materials. The present experiment investigates the difference between intentional and incidental learners as a function of the nature of the orienting task with different stimulus materials. Two orienting tasks were used: (*a*) giving meaningful associations to the stimulus items, and (*b*) matching stimulus items with geometric figures. The association task was more compatible with associative learning than the figure-matching task. There were two sets of learning materials: (*a*) a list of 30 nonsense syllables and (*b*) a list of 30 adjectives. Both intentional and incidental learners performed the orienting tasks. In addition, there were intentional control groups which did not perform the orienting task. The difference between intentional and incidental *Ss* is reduced when both groups perform the orienting task. With nonsense materials, the difference between intentional and incidental learners varies as a function of the nature of the orienting task. When the task is relatively favorable to learning (associations) there is a large difference in favor of intentional learners; with the relatively unfavorable task (figure-matching) there is no difference between the two groups of learners. With meaningful materials, the difference between intentional and incidental learners remains relatively independent of the nature of the orienting task.

References

1. Glaze, J. A. The association value of non-sense syllables. *J. genet. Psychol.*, 1928, **35**, 255–267.
2. Haagen, C. H. Synonymity, vividness, familiarity and association value ratings of 400 pairs of common adjectives. *J. Psychol.*, 1948, **27**, 453–463.
3. Mosteller, F., & Bush, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology*. Cambridge: Addison-Wesley, 1954.
4. Postman, L., & Phillips, L. W. Studies in incidental learning: I. The effects of crowding and isolation. *J. exp. Psychol.*, 1954, **48**, 48–56.
5. Postman, L., Adams, P. A., & Phillips, L. W. Studies in incidental learning: II. The effects of association value and of the method of testing. *J. exp. Psychol.*, 1955, **49**, 1–10.
6. Saltzman, I. J. The orienting task in incidental learning. *Amer. J. Psychol.*, 1953, **64**, 593–598.
7. Tukey, J. W. Comparing individual means in the analysis of variance. *Biometrics*, 1949, **5**, 99–114.

(Received April 23, 1955)[a]

[a] Editor's note.—Publication of this article was delayed for one month in order to print it in the same issue with III and V in the series, which precede and follow it in this issue.

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

STUDIES IN INCIDENTAL LEARNING: V. RECALL FOR ORDER AND ASSOCIATIVE CLUSTERING[1]

LEO POSTMAN, PAULINE AUSTIN ADAMS, AND AUDREY M. BOHM

University of California

Comparisons between intentional and incidental learning have been concerned primarily with the amounts retained under the two conditions of practice. The present paper extends the comparison between the two types of learners to other characteristics of retention which are of significance in serial learning. Two experimental questions will be considered: (*a*) to what extent do intentional learners surpass incidental learners in the retention of *serial order* as well as in the number of items recalled; (*b*) if the items of the series can be grouped into clusters according to a principle, to what degree will intentional and incidental learners differ in the amount of such *associative clustering?*

Recall for serial order.—If the instructions to the intentional learners merely call for memorization of the material and make no explicit reference to serial order, these *S*s may be expected to direct their effort primarily to the differentiation and fixation of the individual items. Their set, however, should generalize to serial order to some extent from the beginning of practice (cf. **6**). Incidental learners are, of course, assumed to be free of any set to learn. Under these conditions, intentional and incidental learners should diverge more in the free recall of the items than in the reconstruction of the serial order. The intentional learners, however, should surpass the incidental learners in both respects.

When the instructions to the intentional learners are extended to include the order of the items, we should expect these *S*s to show an improvement in the reconstruction of serial order, accompanied by a decrease in the amount of free recall. The addition of instructions for order to the intentional *S*s should, therefore, (*a*) increase the difference between intentional and incidental learners with respect to the reconstruction of serial order, and (*b*) decrease the difference between the two groups on a test of free recall. These predictions should hold regardless of whether or not the intentional *S*s perform the orienting task assigned to the incidental *S*s. These hypotheses are tested in Exp. I.

Associative clustering.—A set to group the items may be established not only by explicit instructions but also by manipulation of the stimulus materials. When the stimulus items fall into a number of distinct categories, *S*s may develop a set to group the items in terms of these categories. To the extent that exposure to the stimulus materials establishes such a set in both intentional and incidental learners, both groups should show a tendency toward the associative clustering of related items. The difference between the two groups of learners should, therefore, be greater with respect to the amount recalled than with respect to the degree of associative clustering. This hypothesis is tested in Exp. II. The predictions for Exp. I and Exp. II agree in expecting differences in amount of

[1] This research was facilitated by a grant from the Behavioral Sciences Division of the Ford Foundation.

recall to appear more readily than differences in the retention of systematic characteristics of the series.

Experiment I

Method

Learning materials.—The learning materials consisted of a list of 20 nonsense syllables, with association values of 40% to 46.67% by Glaze's norms (1). Three different serial orders of the list were used. Each serial order was used equally often under all conditions of the experiment.

Incidental learning.—The orienting task used to expose incidental *Ss* to the stimulus list consisted of matching the nonsense syllables with geometric designs. For each syllable, *Ss* were instructed to select one of the following geometric designs as best fitting the stimulus item: square, triangle, circle, cross, arrow, diamond, and wavy line (sine curve). The instructions called for quick "intuitive" judgments without regard to the associations suggested by the stimulus items. Each syllable was exposed for 4 sec. on a 2 × 2-in. slide. The height of the letters on the screen was 1.25 in. The *Ss* had 6 sec. in which to draw the figure they had chosen in the appropriately numbered space of a record blank. Throughout the training procedure, *Ss* were provided with a cardboard strip to cover up all spaces of the record blank except the one in which they were writing. The series was presented four times, each time in the same order. The time interval between successive series was 25 sec. Two sets of variations of the geometric designs were used, one set of variations being used on the first and third trials, the other set on the second and fourth trials. These variations were used in order to make the repeated presentation of the list plausible.

Intentional learning.—There were four groups of intentional learners. Two groups—the Intentional Task Groups—performed the same orienting task as the Incidental Group and, in addition, were instructed to learn the nonsense syllables. Two other groups—the Intentional Control Groups—did not perform the orienting task and were given learning instructions only. One of the Intentional Control Groups and one of the Intentional Task Groups were instructed to learn the syllables in serial order; the instructions to the other two Intentional Groups made no reference to serial order. The rate and method of presenting the syllables were the same for the Intentional Groups as for the Incidental Group, except for the omission of the orienting task in the case of the Intentional Control Groups. For all Intentional Groups, the learning instructions were repeated before each presentation of the series.

Retention tests.—For purposes of the retention test, each of the learning groups was subdivided into two test groups. One of the test groups was given a test of free recall; the other test group received a test of reconstruction. The *Ss* given a free recall test were instructed to write down as many syllables as they could remember, regardless of the order of presentation. For the reconstruction test, *Ss* were presented with the list of syllables in alphabetical order and instructed to arrange them in the order of presentation. The three series used in the experiment were such that the correlation between the correct order and the alphabetical order was very close to zero. Both retention tests had a time limit of 5 min.

With five conditions of learning—one incidental and four intentional—and two types of retention tests, the total experimental design comprised 10 groups. There were 21 *Ss* in each of the groups.

Subjects.—The *Ss* were 210 undergraduate students at the University of California. They did not know the purpose of the experiment. The *Ss* were run in small groups averaging about three in each. The groups were assigned to the experimental conditions in rotation as they appeared in the laboratory, with correction for equal *N*'s at the end. Twenty-one *Ss* were discarded and replaced for the following reasons: 3*Ss* failed to follow instructions during the performance of the orienting task; 4 *Ss* failed to follow instructions on the test of free recall, e.g., by recording the geometric designs rather than the syllables; 11 *Ss* failed to follow instructions on the reconstruction test, e.g., by not using the total series of items in the reconstruction of the serial order. Only two incidental *Ss* had to be discarded because they tried to memorize the syllables.

Results

Free recall.—Table 1 shows the average numbers of items correctly reproduced on the test of free recall. We note the following general trends: (*a*) all Intentional Groups show a higher level of recall than the Incidental Group; (*b*) in agreement with previous findings (**5, 10**), the Intentional Task Groups have lower recall scores than the Intentional Control Groups; (*c*) under both Intentional

TABLE 1

MEAN RETENTION SCORES

Groups	Free Recall		Reconstruction	
	Mean	*SD*	Mean	*SD*
Incidental	4.48	3.36	.26	.30
Intentional Control				
No order instructions	12.95	1.53	.67	.23
Order instructions	10.14	1.69	.86	.17
Intentional Task				
No order instructions	7.81	2.61	.38	.31
Order instructions	5.38	3.15	.54	.22

Task and Intentional Control conditions, the group which has instructions for order recalls fewer items than the group which has general learning instructions only. Analysis of variance shows the difference among conditions to be highly significant ($F = 16.74$, $df = 4$ and 100, $P < .01$). Tukey's gap test (**13**) was used to determine the significant breaks between individual conditions. All breaks between successive groups are significant at either the .05 or the .01 level, with one notable exception: the Intentional Task Group which also has instructions for order does not exceed the Incidental Group by a significant amount.[2] Both instructions for order and the imposition of the orienting task reduce the amount of retention. When these detrimental influences summate, intentional learners no longer surpass the incidental learners.

To assess the interaction between orienting task and instructions for order, a two-way analysis of variance was performed on the scores of the Intentional Groups. Both orienting task and instructions are highly significant sources of variance. For instructions, $F = 9.02$ ($df = 1$ and 80, $P < .01$); the comparison of the Task Conditions and Control Conditions yields an F of 32.25 ($df = 1$ and 80, $P < .01$). However, the interaction of orienting task and instructions is clearly not significant ($F = .77$). The instructions for serial order have parallel detrimental effects on recall whether or not the intentional learners perform the orienting task.

Reconstruction of serial order.— Rank-order correlations between the correct sequence and the reconstructed sequence of items were used to measure *Ss'* retention of serial order. Table 1 shows the mean correlations for the various groups. The various groups differ considerably in the accuracy of recall for serial order: (*a*) all Intentional Groups surpass the Incidental Group; (*b*) the Intentional Control Groups surpass the Intentional Task Groups; and (*c*) for both the Intentional Task Groups and the Intentional Control Groups, instructions for order raise the accuracy score.

Since the distributions of correlations tend to be skewed, nonparametric tests were used in determining the significance of differences among the groups. A ranked one-way analysis of variance (**4**), which makes no assumptions about the distributions of scores, shows the variation among groups to be highly significant ($H = 43.68$, $df = 4$, $P < .01$). Wilcoxon's test for unpaired replicates (**14**) was used to evaluate the significance of differences between individual groups. All Intentional Groups are significantly different from the Incidental Group at the .01 level, with one important exception: the Intentional Task Group which does not receive instructions for order does not differ significantly from the Inci-

[2] The present orienting task has been used in an earlier study. With one trial no significant difference was found between incidental learners and intentional learners performing the orienting task (**5**). With four practice trials, the Intentional Task Group without instructions for order significantly surpasses the Incidental Group.

dental Group. The same conditions of training, however, had produced a substantial difference in free recall. The Intentional Groups are significantly different from each other at the .01 level, with the following two exceptions. The difference between the two Intentional Task Groups—the one which did and the one which did not receive order instructions—is of marginal significance ($P = .07$ by the two-tail test). The performance of the orienting task reduces the effectiveness of the intentional *Ss*, so that specific instructions for order have only limited effects. Finally, the difference between the Intentional Task Group with order instructions and the Intentional Control Group without order instructions is not significant ($P = .10$). Performance of the orienting task and instructions for order have opposite effects on the accuracy of reconstruction. Simultaneous variation of both these factors does not leave a significant difference.

Comparison of recall and reconstruction.—To compare the results of the test of free recall and the test of reconstruction, we shall use the index, $100 \frac{\text{(Intentional–Incidental)}}{\text{Intentional}}$. This index, which may be designated as the percentage "intention effect," measures the relative superiority of intentional over incidental learners independently of the absolute level of the retention scores. Table 2 shows the percentages of intention effect as a function of the conditions of intentional learning. The variations in the index support the following conclusions: (*a*) imposition of the orienting task on the intentional *Ss* lowers their general level of performance and hence reduces the difference between intentional and incidental learners; (*b*) in the absence of instructions for order the intention effect is greater in free recall than in reconstruction; (*c*) the addition of instructions for order raises the intention effect in reconstruction but at the same time lowers the effect in free recall; (*d*) the relationship between the specific nature of the instructions and the intention effect is parallel for the Intentional Control Groups and Intentional Task Groups although the variations tend to be more pronounced in the latter groups. Thus, the hypotheses which Exp. I was designed to test are confirmed.

TABLE 2

RELATIVE SUPERIORITY OF DIFFERENT GROUPS OF INTENTIONAL LEARNERS OVER INCIDENTAL LEARNERS WITH RESPECT TO NUMBER OF ITEMS RECALLED AND RECONSTRUCTION OF SERIAL ORDER

(Entries are 100 [Intentional–Incidental]/ Intentional)

Groups	Number Recalled	Reconstruction
Intentional Control		
No order instructions	65.44	50.55
Order instructions	55.87	61.06
Intentional Task		
No order instructions	30.49	19.16
Order instructions	16.81	39.30

EXPERIMENT II

Experiment II extends a design previously used by Jenkins and Russell (**2**). These investigators presented to their *Ss* a list of words consisting of stimulus-response pairs from the Kent-Rosanoff list (**3**). The *Ss* were instructed to learn as many of the items as possible. The words were presented in random order, but *Ss* showed a strong and highly significant tendency to recall the members of associative pairs together. Recalls in the stimulus-response sequence occurred more frequently than recalls in the response-stimulus sequence, although stimulus words pre-

ceded and followed response words in the list about equally often. The present study compares intentional and incidental learners with respect to the degree of such associative clustering.

Method

Stimulus list.—The stimulus list and order of presentation were exactly the same as those used by Jenkins and Russell (2). The list consists of 48 words—24 pairs from the Kent-Rosanoff tables. These pairs were selected from the 40 most frequent pairs according to Schellenberg's norms for college students (11). Overlapping stimulus and response pairs were avoided. The members of the pairs were distributed at random through the list, with the restriction that members of a pair did not appear contiguously in the stimulus-response order. For 11 of the pairs the stimulus member preceded the response member in the list; for 13 of the pairs the order of appearance was reversed.

Incidental learning.—The incidental *Ss* were given the orienting task of rating each stimulus word with respect to the frequency with which they used it in their daily speech. A 7-point rating scale was used, with 1 standing for extremely infrequent, 7 for extremely frequent, and the other numbers indicating intermediate degrees of frequency. The list, read aloud by *E*, was presented once. After each word *Ss* had 4 sec. in which to record their ratings in the appropriately numbered space of a record blank. The incidental *Ss* will be referred to as Incidental Raters.

Intentional learning.—There were two groups of intentional learners, viz., (*a*) Intentional Raters and (*b*) Intentional Controls. The Intentional Raters performed the same orienting task as the incidental *Ss*, and in addition were instructed to remember as many of the words as possible, regardless of the order of presentation. The Intentional Controls did not perform the orienting task and were given learning instructions only.

Recall tests.—Retention was tested by the method of free recall. The *Ss* were given 5 min. in which to write down as many of the words as they could remember. The *Ss* arranged the words in columns on the page, so that the order in which the words were reproduced could be determined. In order to compare the degrees of clustering which would occur spontaneously and under explicit instructions to reproduce the items in pairs, two sets of instructions were used at the time of recall: (*a*) *Ss* were simply instructed to write down as many words as they could remember, in any order that occurred to them; (*b*) *Ss* were requested to write down as many pairs as possible. In addition, they were, of course, to write down all the single words they could recall. For these *Ss*, special test sheets were used which were marked off with columns of lines headed "pairs" and "single words," so that *Ss* could place the words in the appropriate categories as they occurred to them. This method of recording made it possible to identify new combinations of words which were perceived as pairs by the *Ss*.

For each of the three types of training, a group of 25 *Ss* was tested under conditions of spontaneous clustering, and another group of 25 *Ss* under conditions of directed clustering. The three groups tested under conditions of spontaneous clustering will be discussed under the heading of Exp. IIA; the three groups tested under conditions of directed clustering comprise Exp. IIB.

Subjects.—The *Ss* were 150 undergraduate students at the University of California. They did not know the purpose of the experiment. The *Ss* were run in small groups, averaging about three in each. The groups were assigned to the experimental conditions in rotation as they appeared in the laboratory, with correction for equal *N*'s at the end. None of the incidental *Ss* guessed the purpose of the experiment. Thus, no *Ss* had to be discarded.

Results

Amount retained.—Table 3 shows the total recall scores of the various groups. The overall level of recall is highest for the Intentional Controls, lowest for the Incidental Raters, and intermediate for the Intentional Raters. This rank order is in agreement with the general finding that performance of the orienting task lowers the learning efficiency of the intentional *Ss*. The range of differences is greater under conditions of spontaneous than directed clustering. For Exp. IIA, analysis of variance shows the differences in total recall scores to be significant ($F = 5.03$, $df = 2$ and 72, $P < .01$). The significant break occurs between the incidental and two intentional condi-

TABLE 3

MEAN NUMBERS OF SINGLE AND PAIRED ITEMS RECALLED: EXP. II

Recall Score	Experiment IIA						Experiment IIB					
	Intentional Controls		Intentional Raters		Incidental Raters		Intentional Controls		Intentional Raters		Incidental Raters	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Single items	12.36	5.92	11.32	4.38	10.00	3.54	5.92	3.00	4.72	2.20	6.64	3.96
Forward pairs	4.16	2.05	3.72	2.01	3.52	2.02	5.72	2.65	5.64	2.17	4.76	2.25
Backward pairs	3.20	2.19	3.52	2.45	2.60	1.81	5.68	3.15	4.64	1.85	4.44	2.21
Total pairs	7.36	3.65	7.24	3.74	6.12	3.37	11.40	4.40	10.28	2.92	9.20	3.52
Total items	27.08	5.90	25.80	5.27	22.24	5.22	28.72	6.81	25.28	4.79	25.04	6.77
Per cent paired items	52.83	21.23	54.36	19.47	52.06	20.04	76.99	16.31	79.87	12.01	77.93	17.59

tions. In Exp. IIB, the difference between Incidental Raters and Intentional Raters all but disappears while Intentional Controls retain some superiority. The difference among the conditions, however, falls short of significance ($F = 2.65$, $P > .05$). The convergence of Intentional Raters and Incidental Raters in Exp. IIB is due primarily to an increase in the scores of the incidental group. The explicit search for pairs provides *Ss* with contextual cues. Previous experimentation has shown that a facilitative context at recall reduced the difference between intentional and incidental learners (**7**). In the present case this interpretation, however, must remain tentative since the interaction between test instructions and conditions of training is not significant ($F = 1.01$).

Associative clustering at recall.— Table 3 shows the average numbers of single items, forward pairs (stimulus-response sequence) and backward pairs (response-stimulus sequence) reproduced by the various groups.[3] The percentages that paired items are of the total recall scores are also shown. In agreement with Jenkins and Russell, we find that pairs are reproduced more frequently in forward than in backward order. The difference in favor of forward pairs is significant in Exp. IIA ($F = 7.45$, $df = 1$ and 72, $P < .01$) but not in Exp. IIB ($F = 1.49$). Under conditions of directed clustering the number of pairs is maximized, and the difference between forward and backward pairs, which appears clearly under the more spontaneous conditions of Exp. IIA, is obscured. Preference for forward pairs is not systematically related to the conditions of training in either half of the experiment.

There is only little variation among the groups in the percentages of paired items. In testing the significance of differences in the percentages of paired items, the arc-sine transformation was used (**12,** pp. 448–450). The differences among the groups are not significant in either Exp. IIA ($F = .09$) or Exp. IIB ($F = 1.44$). Degree of associative clustering is, therefore, independent of the conditions of training. Explicit instructions to reproduce the items in pairs result in a considerable increase in the degree of

[3] In Exp. IIB, pairs other than those from the Kent-Rosanoff tables were given only infrequently. Thus *Ss'* categorization of the items confirms the associative norms. The new pairs which did appear were not counted as paired reproductions.

TABLE 4

MEAN NUMBERS OF PAIRS OF RELATIVELY HIGH AND LOW ASSOCIATIVE STRENGTH REPRODUCED

Groups	Experiment IIA				Experiment IIB			
	High Strength		Low Strength		High Strength		Low Strength	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Intentional Controls	3.68	1.89	3.68	2.29	5.40	2.32	6.00	1.93
Intentional Raters	4.32	2.01	2.92	2.06	5.84	1.54	4.44	1.98
Incidental Raters	3.92	1.89	2.20	1.92	5.16	2.33	4.04	1.86

associative clustering, but the increases are of approximately equal magnitude for the three groups. For each of the three groups, the difference in the percentage of paired items between Exp. IIA and IIB is significant at the .01 level. However, the interaction, Conditions of Training × Test Instructions, is not significant ($F = .66$), so that the increase in degree of clustering does not vary reliably from group to group.

Clustering as a function of associative strength.—Although the three groups show the same degree of associative clustering, they differ with respect to the specific pairs reproduced. The 24 pairs in the list were dichotomized into two sets: 12 pairs of relatively high associative strength and 12 pairs of relatively low associative strength. For each pair, degree of associative strength was defined by the frequency of the stimulus-response sequence in the association tables. The recent Minnesota revision of the Kent-Rosanoff norms was used in determining the strength of associative connection (**9**).[4] Table 4 shows the average number of pairs of high and low associative strength recalled by the different groups. The relative frequency of the two types of pairs differs from group to group. Intentional Controls recall pairs of relatively high and low associative strength in about equal proportions. For both Intentional and Incidental Raters, the number of pairs of high strength exceeds the number of pairs of low strength. The difference between the numbers of the two types of pairs was determined for each *S*, and the distributions of difference scores were subjected to an analysis of variance. The variation among the groups was significant at the .01 level in both Exp. IIA ($F = 6.12$, $df = 2$ and 72) and Exp. IIB ($F = 5.86$), with the significant breaks occurring between the Intentional Controls and the two groups of raters. The two halves of the experiment do not differ significantly with respect to the distribution of difference scores. The finding that incidental *S*s favor pairs of high strength is in agreement with the results of an earlier study which showed that intentional learners surpass incidental learners in the recall of items of low but not of high association value (**7**). The present results indicate that intentional *S*s performing the orienting task show

[4] The revised Minnesota norms became available after the publication of the study by Jenkins and Russell. The new norms were used in preference to Schellenberg's 1930 norms in estimating the associative strength of the stimulus-response pairs.

similar selectivity for materials of high associative strength.

Discussion

"Intent" to learn cannot be characterized in all-or-none terms but rather varies in the degree to which it is focused on different characteristics of the learning materials. Intentional learners will surpass incidental learners only insofar as their set encompasses dimensions of the stimulus which are relevant to performance on the test of retention. The intentional *S* is to some extent an incidental learner with respect to characteristics on which his set is not centrally focused. General instructions to learn a series of items direct the intentional *S* primarily to the differentiation and fixation of individual items and to a lesser degree to serial order. Hence, in the absence of explicit instructions regarding order, the two groups differ less on the test of reconstruction than on the test of free recall. When the instructions explicitly refer to serial order, the difference in the reconstruction scores is increased, but only at the expense of a reduction of the difference on the test of free recall (Exp. I).

Selective sets can be established not only by instructions but also by the recognition of recurrent systematic features of the learning material. Once *S* has identified some rule governing the sequence of stimulus items, he is ready to discriminate and categorize new instances in accordance with that rule. Such stimulus-produced sets may be established in incidental as well as intentional learners. The more readily identifiable the systematic characteristics of the series, the less should intentional and incidental learners differ in the recall of such characteristics. The results of Exp. II illustrate this point. The strong associative connection between the stimulus and response members of the word pairs was readily discriminated by both intentional and incidental *S*s. Hence there were no differences in the percentage of paired items reproduced by the two groups of learners. It is noteworthy, however, that both incidental *S*s and intentional *S*s performing the orienting task showed selective retention for the pairs of highest associative strength. Absence of instructions to learn and interference from the orienting task make *S*s selectively sensitive to those items which conform most clearly to the principle of the series.

The two experiments support the conclusion that the degree of difference between intentional and incidental learners depends on (*a*) the specific stimulus characteristics for which retention is tested, and (*b*) the relative readiness of the two groups to respond differentially to these characteristics.

Summary

Intentional and incidental learners were compared with respect to (*a*) retention of serial order and (*b*) associative clustering of related members of a learning series. In Exp. I, the learning materials consisted of a list of 20 nonsense syllables. Half of the intentional *S*s were given instructions to remember the serial order of the items while the other half were given general learning instructions only. Retention was tested either by free recall or reconstruction. In the absence of specific instructions for order, intentional learners surpassed incidental learners more on the test of free recall than in the reconstruction of serial order. Explicit instructions for order raised the superiority of the intentional *S*s on the test of reconstruction but at the same time reduced the difference in free recall.

In Exp. II the stimulus materials were 24 stimulus-response pairs from the Kent-Rosanoff association tables. Stimulus and response members were scattered at random through the list. Half the *S*s were given a conventional test of free recall whereas the other half were instructed to reproduce as many stimulus-response pairs as possible. The intentional *S*s recalled a larger number of items than the incidental *S*s. The differences in the amount of recall were significant when clustering was spontaneous but not when *S*s were instructed to reproduce the items in pairs. The percentages of items reproduced in pairs were similar for intentional and incidental learners, i.e., degree of associative clustering was independent of the conditions of training.

References

1. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–267.
2. Jenkins, J. J., & Russell, W. A. Associative clustering during recall. *J. abnorm. soc. Psychol.*, 1952, **47**, 818–821.
3. Kent, G. H., & Rosanoff, A. J. A study of association in insanity. *Amer. J. Insanity*, 1910, **67**, 37–96, 317–390.
4. Kruskal, W. H., & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Ass.*, 1952, **47**, 583–621.
5. Postman, L., & Adams, P. A. Studies in incidental learning: IV. The interaction of orienting tasks and stimulus materials. *J. exp. Psychol.*, 1956, **51**, 329–333.
6. Postman, L., & Senders, V. L. Incidental learning and generality of set. *J. exp. Psychol.*, 1946, **36**, 153–165.
7. Postman, L., Adams, P. A., & Phillips, L. W. Studies in incidental learning: II. The effects of association value and the method of testing. *J. exp. Psychol.*, 1955, **49**, 1–10.
8. Postman, L., & Adams, P. A., Studies in incidental learning: III. Interserial interference. *J. exp. Psychol.*, 1956, **51**, 323–328.
9. Russell, W. A., & Jenkins, J. J. The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff word association test. In *Studies on the role of language in behavior*, Univ. of Minnesota, 1954. Tech. rep. No. 11.
10. Saltzman, I. J. The orienting task in incidental learning. *Amer. J. Psychol.*, 1953, **64**, 593–598.
11. Schellenberg, P. E. A group free association test for college students. Unpublished doctor's dissertation, Univ. of Minnesota, 1930.
12. Snedecor, G. W. *Statistical methods.* Ames: Iowa State Coll. Press, 1946.
13. Tukey, J. W. Comparing individual means in the analysis of variance. *Biometrics*, 1949, **5**, 99–114.
14. Wilcoxon, F. *Some rapid approximate statistical procedures.* New York: American Cyanamid Co., 1949.

(Received August 8, 1955)[a]

[a] Editor's note.—Publication of this article was advanced three months in order to print it in the same issue with III and IV in the series, which precede it in this issue.

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

PERCEPTUAL THRESHOLDS OF CURVILINEARITY AND ANGULARITY AS FUNCTIONS OF LINE LENGTH

LOUIS DELLA VALLE,[1] T. G. ANDREWS, AND SHERMAN ROSS

University of Maryland

This experiment deals with the perception of horizontal lines and the minimum change in their direction which is detectable, when the change is either a regular curve or an angle. We are interested in this aspect of foveal visual acuity under photopic conditions and as a function of the basic line length.

Takagi (**4**) in 1926 was one of the first investigators interested in perceptual parameters of curved lines. His *S*s judged the lengths of a variety of arcs and straight lines. Takagi found that arc lengths could be estimated almost as accurately as a straight line. The estimation of the arc length, however, was influenced by its curvature. The greater the curvature, the greater was the overestimation of the arc.

A second purpose of the present investigation was to obtain threshold measures for the perception of simple curvature. It is to be noted that this type of visual acuity has not been subjected to detailed experimental study as have the more standard aspects of the minimum perceptible (size), minimum separable (gap), and minimum distinguishable (vernier or discontinuity) (**5**). A previous study from the Maryland laboratory reported on the effects of surround variation on the perception of curvature (**3**). The results of this experiment indicated that thresholds for the curvature of stimulus lines tended to be higher when the stimulus was placed in a surrounding form. While the degree of curvature was varied in this experiment, the chord length of the stimulus line was kept constant at 3 cm.

Another study related to the present one is the work of Gibson on adaptation, after effect, and contrast in the perception of curved lines (**2**). In his experiment Gibson did not, however, use precise degrees of curvature, nor was the variable of chord length manipulated.

In this experiment we have tested two hypotheses: (*a*) as base or chord length is increased, thresholds for curvature and for angularity of lines are decreased; and (*b*) thresholds for angularity are lower than for curvilinearity of lines. The first hypothesis is deducible from the increase in the number of retinal elements involved. The second hypothesis is less compelling. It is generated from the judgment that curved lines, from their more regularly changing contour, possess greater *prägnanz* than do angular arrangements. Thus curved lines may require a greater change in direction in order to be perceived as "off-straight."

PROCEDURE

Subjects.—The *S*s were 10 male university students ranging in age from 18 to 28 years. Each *S* had a normal Snellen index of 20/20 uncorrected for both eyes.

Materials.—The curved and angular stimulus lines were drawn in India ink with a Chavos pen and centered on white cards 14 in. square. The lines were approximately .4 mm. wide. For the curved lines exact and known degrees of

[1] Now at the University of Bologna, Italy.

curvature were produced by using brass optical templates, rated in diopters, procured from the American Optical Company, Southbridge, Massachusetts.[2] The lengths of the stimulus lines used were 30, 40, 50, 70, 90, and 110 mm. as measured by the chord of the curves and the base length of the obtuse angles.

The range of curvatures selected was based upon a previous study (3) in which it was found that a 1.5 diopter curve with a chord length of 30 mm. was correctly identified as curved 100 per cent of the time. This value of 1.5 diopters was therefore selected as the upper limit of curvature for the present experiment. The curves used have the following diopter values: 0, .37, .75, 1.12, and 1.5. The arc radius in meters equals .53 divided by the template diopter value. The above arcs in terms of the radii of their curvatures in centimeters are: plano (infinity), 143.24, 70.67, 47.32, and 35.33, respectively.

The angular lines were determined by the specific curved lines selected. It was necessary to equate the size of the area of a curved line to an angle of the same base length. This was done in order to make certain that discriminations of both types of stimuli, for a given length of line, were not being made on the basis of differential size of the retinal area stimulated. By area of the stimulus line is meant the area *above* the curved line or angular line to a horizontal line perpendicular to the sagitta. The sagitta is the perpendicular line from the center of the chord to the arc or from the center of the base of the angular line to the apex. Detailed mathematical analyses revealed that the area above a curved line approximates the area above an angular one of equivalent base length when the sagitta of the angular line is one-half that of the curve.[3] The stimulus lines were therefore made so that the bases of the angular lines were equal to the chords of the curved lines, and the sagitta of the angular lines were one-half the sagitta of the "equivalent" curved lines. Following this procedure we found that the angular line could not be drawn accurately for the 30-mm. base length. The sagitta of the curves were computed as $s = y^2/2r$, where y is half the chord length and r is the radius of the arc as computed from the template diopter value. There were 30 stimulus cards for the curved lines and 25 stimulus cards for the angular lines.

Apparatus.—The apparatus for presenting the stimulus cards was a rectangular Masonite viewing screen 42 in. long and 30 in. high with a center aperture 13 in. square behind which the stimulus cards were exposed. The *S* was seated 10.5 ft. from the apparatus in order to keep the stimulus within the foveal limit of 2°. The experimental room was darkened, and a Kodaslide projector used to illuminate the stimulus lines on the white cards at approximately 4 ft.-c. The light beam from this source was passed through an aperture with uneven and sharply serrated edges in order to remove any even shadow contours and reference edges for the *S* to use as cues in making his judgments.

Stimulus presentation.—After *S* adapted to the level of illumination provided by the stimulus screen, he was told that he would be shown cards on which would be horizontal black lines. Each curve and angle was presented with its concavity downward. The *S*'s task was to report whether the line in each case was "straight" or "not straight." The curved and angular lines were randomly presented in a mixed series, and each stimulus was presented 10 times. No two *S*s received the same order of stimuli. There was no time limit set for the discriminations, and most *S*s responded within approximately 4 to 5 sec. Three sessions of one hour each were required for each *S* to complete the 550 discriminations required. A 5-min. rest period was given after the first half-hour of each session.

Results

The judgments of the *S*s were recorded and treated in terms of the method of single stimuli. Thresholds were determined at the 75% point by linear interpolation on the psychometric functions of the "not straight" category. Table 1 presents for each line length the limens in diopters, radius of arc, sagitta, and adjusted sagitta separately for the curved lines and for the angular lines. These results require certain interpretations because of their geometric characteristics. Diopter values are proportional to degree of curvature, and it can be seen that thresholds of curvature in terms of this measure decrease with increased length of line

[2] Since this work was done the writers have learned of a new method for producing accurate arcs of very large radii (1).

[3] For this geometric analysis we are indebted to Dr. Guilford L. Spencer of the Department of Mathematics, University of Maryland. We can furnish copies of this analysis upon individual requests.

TABLE 1

THRESHOLDS OBTAINED FOR CURVED AND ANGULAR LINES FOR SIX LINE LENGTHS

Line Length in mm.	Curved Lines				Angular Lines			
	Threshold in Diopters	Threshold in cm. Radius*	Threshold in mm. Sagitta**	Threshold in mm. Adjusted Sagitta***	Threshold in Equivalent Diopters	Threshold in cm. Equivalent Radius*	Threshold in mm. Sagitta**	Threshold in mm. Adjusted Sagitta***
	I-A	II-A	III-A	IV-A	I-B	II-B	III-B	IV-B
30	1.33	39.85	.28	.28	—	—	—	—
40	1.08	49.07	.41	.23	1.10	48.18	.42	.23
50	.81	65.43	.48	.17	1.06	50.00	.63	.23
70	.54	98.15	.62	.12	.65	81.54	.75	.14
90	.54	98.15	1.03	.12	.66	80.30	1.26	.14
100	.37	143.24	1.06	.08	.64	82.81	1.83	.14

* Radius in cm. = $53/D$.
** Sagitta in mm. = $(.5L^2)/20\,r$.
*** See text.

(Columns I-A and I-B). The diopter value for an angular line, however, is based here on the construction of angles subtending equivalent areas to the standard curves used and is, of course, an unusual measure of "off-straightness." The radius of the arc for the threshold degrees of curvature were computed and are shown in Columns II-A and II-B. These values present an increasing function with length of line as shown in Fig. 1. It is again to be noted that radius of equivalent arc is an unorthodox measure for "off-straightness" of an angular line.

As a measure of acuity in the discrimination of lines and curves it would be useful to measure a particular distance for the "off-straightness." For this purpose we have used the sagitta. The sagitta, however, suffers as an experimental dependent variable in the present case because it is not independent of the line length, as shown by its formula which we presented earlier. In order to obtain an estimate of some experimentally useful value for acuity in this case we have taken the sagitta in each case for the 30-mm. segment of each curve and line. The values of these "adjusted" sagittae are presented in Columns IV-A and IV-B of Table 1. These adjusted sagitta demonstrate a decreasing function of line length. Operating with the only sagitta directly obtained from the data, .28 mm., it is found that this threshold for "off-straightness" of a curved line 30 mm. long represents 1.8 sec. of visual angle, which compares very favorably with estimates of visual acuity reported for other types of stimulus conditions (5). It is expected that acuity would be quite high for perception of curvature.

By all the available estimates, perception of curvature is apparently superior for longer segments of curves. These data, however, do not give a statistical test of significance, and more data are required for a rigorous determination of the exact functional relationship that obtains between line length and discrimination of curvature.

In comparing the columns of Table 1 for curved lines and for angular lines, and as noted in the graphic comparison in Fig. 1, sensitivity to "off-straightness" was consistently

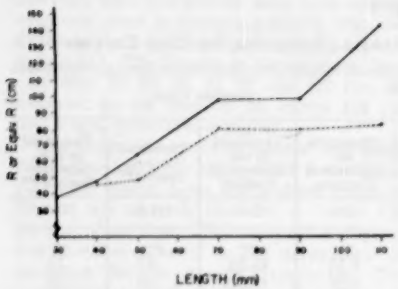

FIG. 1. Showing the mean threshold values (radius and equivalent radius) for the curved and angular stimulus line as a function of chord or base length. The results for the curved stimuli are shown by the continuous line, and those for the angular stimuli for the dashed line.

greater in our sample for curves than for angles. The data do not present a statistically significant difference in trend, and more data are required for a sensitive test of the second hypothesis of the study. The result obtained here, however, represents an interesting reversal of our second hypothesis.

More intensive study of these functions was made on one experimental *S*, J. A. W. Forty judgments for each stimulus for each line length were obtained; a total of 2,200 judgments. This *S* demonstrated the same trends as the group data with respect to thresholds decreasing with line length and thresholds for curves being consistently lower than for straight lines. He consistently produced lower thresholds than the average results presented in Table 1. His threshold sagitta for a curved line of 30 mm. chord was .25 mm., which gives an estimated visual angle of 1.61 sec., showing a rather sensitive mechanism.

Although we place less reliance on it for the reasons given above, the average threshold adjusted sagitta for a curved line of 110 mm. of chord length was estimated as .08 mm., giving an estimated visual angle of .52 sec. There are essentially no straight lines in the natural environment of man except the ones he attempts to construct. The human visual system seems nevertheless to be exquisitely sensitive to deviations from a straight line.

SUMMARY

This experiment was designed to test two hypotheses: (*a*) the thresholds for curvature and for angularity of lines decrease as base length or chord length increase, and (*b*) thresholds for angularity are lower than for curvilinearity of lines.

The stimuli used were lines of regular curvature with the following diopter values: 0, .37, .75, 1.12, and 1.5. Precise degrees of curvature were produced by using brass optical templates. The chord lengths of the curved lines and the base lengths of the angular lines were 30, 40, 50, 70, 90, and 110 mm. The obtuse-angle stimuli were produced so that the area above the curved line approximated the area above the angular one of similar base length. The stimulus lines were approximately .64 mm. wide.

Ten *S*s were used. The stimuli were presented at a distance of 10.5 ft. at a level of illumination of approximately 4 ft.-c. *S* judged whether or not the stimulus presented was "straight" or "not-straight." The judgments and results were treated in terms of the method of single stimuli.

Thresholds were computed for the curved and angular lines using a variety of measures: diopters, length of radius or equivalent radius, length of sagitta, and length of "adjusted" sagitta.

From the available estimates of sensitivity it is relatively clear that the perception of curvilinearity and angularity is superior for greater chord and base lengths. Sensitivity to "off-straightness" was consistently greater for curves than for angles. More data are required for the determination of the exact functional relationship between base length and the discrimination of curvature or angle.

One of the interesting findings of this study is in regard to the visual acuity limits for the perception of curvature using the sagitta as the measure. For one *S* the lowest visual angle discriminated was approximately 1.61 sec., showing the sensitivity of this mechanism.

References

1. Anon. Device for drawing arcs of very large radius. *Res. Revs.*, 1955, 21–22. (Published by Office of Naval Research, Department of the Navy, Washington, D. C.)
2. Gibson, J. J. Adaptation, after-effect, and contrast in the perception of curved lines. *J. exp. Psychol.*, 1933, **16**, 1–31.
3. Muller, P. F., Andrews, T. G., & Ross, S. Effects of surround variation on visual perception of curvature. *J. gen. Psychol.* (in press).
4. Takagi, K. On visual estimation of length of various curves. *Jap. J. Psychol.*, 1926, **1**, 476–498.
5. Tufts College, Institute for Applied Psychology. *Handbook of human engineering data* (2nd Ed.) Port Washington, L. I., N. Y.: U. S. Navy Special Devices Center, 1951. (Tech. Rep. No. SDC 199-1-2, 1951, Part III, Ch. 11, sec. 11, 1–10.)

(Received March 21, 1955)

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

SUGAR AS A REWARD FOR HUNGRY AND NONHUNGRY RATS[1]

MONCRIEFF SMITH AND GLENN C. KINNEY[2]

University of Washington

Rats are known to respond positively to sugar, that is to say, given access to a sucrose solution, they will drink large quantities of it (1). Moreover, having consumed it, they recognize it as food in the sense that they reduce their intake of other foods to compensate for the calories in the sugar (3). It is not clear, however, that they recognize it as food in the sense that hungry animals will accept it more readily than nonhungry ones.

If hunger does influence the response, the manner in which the discrimination is made becomes of considerable interest since it might provide a cue as to the nature of rewards in general. There are many possible theories that could account for an increased preference for sugar on the part of hungry rats, but two seem to be of particular interest because they represent opposing approaches to the problem. These are the "avidity" theory and the need-reduction theory. The first was given that name by Katz (4), but a very similar point of view has been proposed by Richter. Stated in an experimentally testable form, it assumes that the existence of a state of need in an organism will alter the organism's reaction to a class of stimuli, and that the stimulus class will include stimuli arising from the needed substance. Thus, Richter, Holt, and Barelare (6) report that rats deficient in Vitamin B_1 show great avidity for it in their initial contact, seeking it out from several different substances offered, and consuming large quantities of it at the first opportunity. Later studies (2, 5) have failed to confirm their observation, but it serves to illustrate the theory.

The opposing point of view has no explicit spokesman, but Hull's need-reduction theory, for example, would expect the reduction of need, or of drive stimuli, that accompanies ingestion of sucrose to reinforce the response to stimuli arising from the solution and thus to increase the strength of response exhibited by hungry animals.

A possible means of choosing between these points of view lies in the study of the time course of the development of the preferential response, if, indeed, one does develop. The avidity theory would seem to require an immediate difference in the reaction of hungry and nonhungry rats, whereas the need-reduction interpretation requires time for the operation of reinforcement. Whether this time is to be measured in seconds, minutes, or hours is not clear, but a matter of five to ten minutes, at least, would seem reasonable.

A study of consummatory activity from the time of initial contact would seem to be the most direct approach to the problem, and the results of such observations will be reported at a later date. Consummatory activity

[1] This research was supported by a grant from the State of Washington, Initiative 171 funds for research in Biology and Medicine.

[2] Now at the Crew Research Laboratory, Air Force Personnel and Training Research Center, Randolph Air Force Base, San Antonio, Texas.

has its drawbacks as a measure of preference, however, particularly when the primary interest is in a comparison of hungry and nonhungry animals. Mechanical factors, such as distention of the stomach, might limit the amount of consumption by the nonhungry rats and produce differences that could not be found in other measures of preference.

In the present study we observed the rate of a bar-pressing response in a Skinner box with sugar solution as a reward. The small quantities of solution that the rats received as rewards could not have filled them, so that the possibility of mechanical factors influencing the expression of preference was minimized.

Method

Subjects.—The experiment used 52 male rats, Wistar strain, 3–4 mo. old at the beginning of the experiment.

Apparatus.—The apparatus consisted of four identical Skinner boxes, with feeders that moved a small spoon through a trough so as to scoop up and present to the rat a small quantity of the trough contents. This feeder will present any powdered food or a drop of any liquid (about .06 cc.), and in this experiment only liquids were presented. The four feeders were controlled independently of one another, so that if the feeder sounds traveled from one box to another, they could not serve as a signal for reward.

Procedure.—Four conditions were considered in the initial test—hungry and nonhungry rats, rewarded for bar-pressing with either sugar solution or water. The *S*s were run in sets of four with each condition being represented in each set. The experiment was run in three cycles, with four sets of rats in the first cycle, five in the second, and four in the third. For each cycle th procedure was the same.

The training and testing schedule is summarized in Table 1. The purpose of the initial (23 hr.) water-deprivation schedule was to ensure that the bar-pressing response could be well established without the use of a food reward. Day 4, with the 4-hr. water deprivation, was intended to serve as a bridge between the training and test conditions.

On test days all *S*s were 4-hr. water deprived to ensure that the nonhungry *S*s would do some responding.

TABLE 1

Training and Testing Schedule

Test Days	*N*	Deprivation in Hours		Reinforcement
		Food	Water	
1–3 (preliminary training)	52	0	23	Water
4 (preliminary training)	52	0	4	Water
5 (first test)				
H-S	13	23	4	20% sucrose
NH-S	13	0	4	20% sucrose
H-W	13	23	4	Water
NH-W	13	0	4	Water
7 (second test)				
H-S	26	23	4	20% sucrose
NH-S	26	0	4	20% sucrose
9 and 11 (3rd and 4th tests)				
H-S	17	23	4	20% sucrose
NH-S	18	9	4	20% sucrose

On the second test (Day 7) the *S*s who had experienced sugar on the first test were continued under the same conditions. Those who had water on the first test were pooled and reformed into two groups of 13 each, with one being given H-S conditions and the other NH-S conditions. On the second and third cycles of the experiment two additional tests (Days 9 and 11) were run with the *S*s serving under the same conditions they experienced on Day 7. One rat, tested on Days 5 and 7 under H-S conditions became ill after the second test, and had to be dropped from the last two tests. His data for the first two tests seemed normal, and were retained.

The 48-hr. period between successive tests was intended to prevent a serious cumulative deprivation from developing in the hungry *S*s. They had never been put on a formal food-deprivation schedule, and undoubtedly would have gotten insufficient food in a 1-hr. eating period. As it was, they had food and water ad lib for 23 hrs. after each test before the deprivation period started for the next test.

Results

Table 2 gives the means of the four conditions on the first test day. Each mean is the average number of bar presses made by a group of 13 *S*s during the 46-min. test period. The *SD*'s are also given for each group, and they differ considerably from one another. However, Bartlett's test shows the assumption of homogeneity

TABLE 2

MEANS AND *SD*'S OF BAR-PRESSING RESPONSES ON FIRST TEST
(N = 13 each group)

Reward	Drive State with 4-hr. Water Deprivation			
	Hungry		Not Hungry	
	Mean	*SD*	Mean	*SD*
20% sugar solution	173.0	50.8	153.2	42.8
Water	86.8	37.0	141.5	72.3

of variance to be tenable (χ^2 = 6.37; 5% level = 7.82).

An analysis of variance of the two-by-two table showed two significant sources of variance, sugar (F = 11.32, df = 1/48, $P < .01$), and the interaction of sugar and hunger (F = 6.56, df = 1/48, $.05 > P > .01$), when tested against the within-groups variance (2754). Inspection of the table shows that the primary effect was on the group that was hungry and rewarded with water (H-W). This group mean differed significantly from each of the others, as shown by the t test, while none of the differences among the other three groups was significant.

Cumulative response curves for the 46-min. test period add little information. The H-S group responded at a fairly constant rate throughout the period. The rates of the other groups fell off during the test, with the H-W group showing the greatest decline. The other three groups stayed close together for the first 30 min. and the means at that time were 115.4, 112.9, 105.0, and 66.9 for the H-S, NH-S, NH-W, and H-W groups, respectively.

The significant interaction in Table 2 indicates some effect of sugar, but it seemed equally effective for both hungry and nonhungry *Ss*. The difference favors the hungry group, but it is well below the 5% level of significance (t = .96). Information gathered on later test days also failed to show a differential effect due to hunger. On the second test day, for example, the 26 *Ss* of the W groups were regrouped to form new H-S and NH-S groups. The mean of the H-S group on this test was 168.2, compared with 167.2 for the NH-S group. The cumulative surves were superimposed throughout the test period.

Table 3 gives comparisons between hungry and nonhungry *Ss* on their first, second, and third sessions with sugar reward. Test 1, for example, pools the data of the H-S group on test day one and the 13 *Ss* who experienced H-S conditions for the first time on the second test (Day 7). Since only two tests were given in the first cycle of the experiment, the number of *Ss* drops on each successive test. This vitiates day-to-day comparisons within a group, but should not bias comparisons between groups on successive tests. Although the difference favors the hungry *Ss* in each case, none of the differences approaches significance.

Cumulative curves showed no evidence of an initial preference for sugar by the hungry *Ss*. On most of the tests the rate of responding was fairly uniform throughout the 46-min. test period, with the NH-S and H-S groups starting out at the same rate. In no case did the early differences between these groups approach significance, but it is worth noting that, at the end of the first 10 minutes of responding, the NH-S group was slightly ahead of the H-S group on two of the three test days. This difference is the opposite of that predicted by the avidity theory.

Discussion

Table 2 is susceptible to several interpretations. One way of looking at it emphasizes the comparison of differences within the columns. That is to say, shifting from a water to a sugar reinforcement produced an increase of 86.2 (significant) for the hungry *S*s and only 11.7 (nonsignificant) for the nonhungry *S*s. This might mean that sugar is reinforcing for hungry *S*s and not for nonhungry ones.

The other way of looking at the table emphasizes a comparison of the rows. For *S*s reinforced with water, hunger produced a drop of 54.7 (significant), while for those reinforced with sugar it produced a rise of 19.8 (nonsignificant). This view suggests that rats like the taste of sucrose, and that hunger has nothing to do with the reaction. The hungry *S*s do not respond to the sucrose solution as they would to just a liquid, nor as they would to a familiar food. Although there is no evidence in this experiment to prove it, it seems likely that the hungry *S*s would have responded at a much higher rate had the reinforcement been powdered Chow.

This second interpretation assumes the thirst drive to be equal for all four groups. On the face of it, this assumption appears reasonable, since all groups were deprived of water for 4 hrs. preceding the test. However, it might be argued that this 4-hr. period made the nonhungry *S*s more thirsty than the hungry *S*s, since the nonhungry *S*s had dry food present and probably ate some during the period. This would imply that the nonhungry *S*s were responding, in part, because of the reinforcing value of the liquid *qua* liquid, and that a direct comparison of the hungry and nonhungry *S*s reinforced with sugar solution is therefore invalid as a test of reaction to sugar itself. We cannot deny that this is the case, although we doubt that the experiment would have come out very differently had the 4-hr. water deprivation been eliminated on test day.

Despite the question raised in the last paragraph, we prefer the second interpretation. The need-reduction theory of preference does not seem to fit the data, since there was little or no increase in response rate for the H-S groups within test periods, and although Table 3 shows an increase from test to test for the hungry *S*s, this increase did not differ significantly from that shown by the nonhungry *S*s. The initial difference between H-S and NH-S groups predicted by the avidity theory did not seem to be present either, although it might have been obscured by a difference in thirst drive.

TABLE 3

Mean Bar-Pressing Responses of Hungry and Nonhungry Groups on Three Successive Tests with Aperiodic Reward of One Drop of 20% Sucrose Solution

Test	Group H-S		Group NH-S		Sdiff.	*t*
	N	Mean	*N*	Mean		
1	26	170.6	26	160.2	13.5	0.77
2	22	201.5	22	171.9	21.5	1.38
3	17	197.0	18	184.6	25.3	0.49

In short, there seems to be little evidence of a need-related preference. We know that such preferences sometimes develop, and it is curious that here they did not. Perhaps the key is in the small amount of reinforcement received during the experiment. Each *S* got less than 1 cc. of sucrose solution per 46-min. test, and this amount of solution would contain less than 200 mg. of sucrose. Spread over a 46-min. test period, such a small amount of sucrose might well not show any need-reducing effects.

The interaction in Table 2, however, indicates some sort of reaction to sugar. There was enough sugar to taste, so that the necessary conditions for the avidity theory were present, and yet there was little evidence of avidity on the part of the hungry *S*s. We suggest that the taste of sugar provides reinforcement that is independent of either the hunger or thirst drives, and that need-related preferences, when they appear, are

attributable to learning, through the mediation of need or drive reduction. The fact that sugar was no more reinforcing than water for the nonhungry *Ss* is not well explained by this theory, and we must assume that the absence of a difference was due to the particular conditions of this experiment. Other studies (**7, 8, 9**) show that sweet-tasting substances are more readily acceptable than water by naive animals, either hungry or satiated, and that these substances will act as reinforcers.

Summary

Fifty-two thirsty rats were trained to press the bar in a Skinner box, with water as a reward. Four groups of 13 each were then formed, and one hungry and one nonhungry group were tested in the Skinner box with a 20% sucrose solution as a reward, while the other two groups were continued on a water reward. All groups were deprived of water for four hours prior to test. The sugar groups were continued under the same conditions for three additional tests, and the water groups were pooled and redivided nto hungry and nonhungry groups for three additional tests on sugar.

A significant interaction between the effects of hunger and sugar on the first test day revealed that the sugar was an effective reinforcer. However, there was no significant difference between the hungry and nonhungry *Ss* on their initial contact with sugar, or on any later test. Although a need-reduction interpretation could not be ruled out, the results were interpreted as indicating that in this situation taste stimuli determined the reinforcing characteristics of sugar, and that need reduction contributed little or nothing to reinforcement.

References

1. Beebe-Center, J. G. Feelings and emotions. In H. Helson (Ed.), *Theoretical foundations of psychology.* Philadelphia: Van Nostrand, 1951. Pp. 254–317.
2. Cohen, J. D. An experimental investigation of thiamine selection by normal and deprived rats. Unpublished doctor's dissertation, Univ. of Washington, 1953.
3. Hausmann, M. F. The behavior of albino rats in choosing foods: II. Differentiation between sugar and saccharin. *J. comp. Psychol.*, 1933, **15,** 419–428.
4. Katz, D. *Animals and men. Studies in comparative psychology.* New York: Longmans, Green, 1937.
5. Luria, Z. Behavioral adjustment to thiamine deficiency in albino rats. *J. comp. physiol. Psychol.*, 1953, **46,** 358–362.
6. Richter, C. P., Holt, L. E., Jr., & Barelare, B., Jr. Vitamin B_1 craving in rats. *Science*, 1937, **86,** 354–355.
7. Sheffield, F. D., & Roby, T. B. Reward value of a non-nutritive sweet taste. *J. comp physiol. Psychol.*, 1950, **43,** 471–481.
8. Sheffield, F. D., Roby, T. B., & Campbell, B. A. Drive reduction versus consummatory behavior as determinants of reinforcement. *J. comp. physiol. Psychol.*, 1954, **47,** 349–354.
9. Young, P. T. Studies of food preference, appetite and dietary habit. I. Running activity and dietary habit of the rat in relation to food preference. *J. comp. Psychol.*, 1944, **37,** 327–370.

(Received April 29, 1955)

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

REACTIVE INHIBITION AS A FUNCTION OF SAME-HAND AND OPPOSITE-HAND INTERTRIAL ACTIVITY [1]

LEWIS E. ALBRIGHT,[2] C. ROBERT BORRESEN, AND MELVIN H. MARX

University of Missouri

In elaborating on his definition of reactive inhibition (I_R), Hull states (**4**, p. 281) that it is probably dependent on an inhibitory substance localized in the responding effectors which is removed by the blood stream during rest. Miller and Dollard (**8**) and Mowrer and Jones (**9**), who support the peripheral hypothesis, conceive of work inhibition as a state of muscular fatigue or strain. Solomon (**13**) also accepts a somewhat modified view of its peripheral locus. Subsequent evidence of bilateral reminiscence (**3, 10**) and bilateral transfer (**7**) of work decrement, however, has indicated that temporary inhibition is more general than this local fatigue theory would lead one to suppose. Other investigators (**1, 12**) have extended the idea of a generalized work decrement still further by specifically suggesting the existence of a central factor, as well as the muscular-fatigue factor in work decrement.

Establishing the neurophysiological locus of work decrement is not crucial to work inhibition theory because this construct has operational validity independent of such physiological specification. Some workers (**3, 13**) have recently pointed out, however, that more definite experimental information concerning the nature of work inhibition might be helpful in clarifying explanations of both bilateral transfer and of bilateral and conventional reminiscence. The present study was an attempt to provide more experimental information about work decrement and the nature and amount of effector activity.

On the basis of the peripheral hypothesis, two specific predictions were advanced: (*a*) Constant use of the same hand, regardless of the variety of tasks performed, results in a depressed performance function similar to that exhibited by *Ss* performing a single task under massed practice conditions. (*b*) The use of the opposite hand to perform an alternate task permits the first hand to dissipate inhibition, resulting in a performance level similar to that attained by *Ss* given distributed practice on a single task.

METHOD

Subjects.—The *Ss* in this study were 169 students from two introductory psychology classes at the University of Missouri. They were assigned to one of four groups representing the four conditions. The method of assignment consisted of arbitrarily dividing each of the two classes in half. Since these two classes met at different hours of the day, *Ss* in the earlier class were asked not to discuss the experiment with their friends in the later section. The conditions for the two groups in the earlier class were similar enough to permit running these two groups simultaneously in the same room. However, a different time schedule for one of the two groups in the later class required that this group be run in a separate room.

Upon examination of the data, it was found that the papers of 16 *Ss* had to be discarded because of failure to follow instructions correctly or because of interruptions due to broken pencils. The discards per group were as follows: 3 in the OH group, 6 in the DP group, 5 in the MP group, and 2 in the SH group.

Learning task.—The learning task was the inverted alphabet-printing task similar to that used by several other investigators (**2, 5, 6, 11, 14**). The *Ss* were given large sheets of blank

[1] We are indebted to Robert A. Goldbeck for his critical reading of the manuscript.

[2] Now at Purdue University.

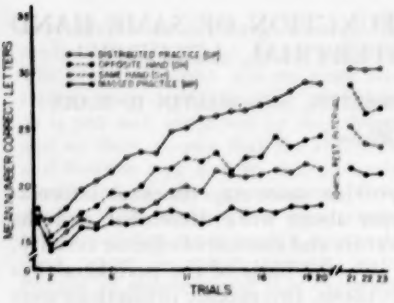

FIG. 1. Mean number of letters correctly printed as a function of practice.

paper on which to record their responses. At the end of each trial they were instructed to draw a short vertical line to the left of the last letter printed. The following trial was then begun to the left of this line. These "unstructured" response sheets prevented the occurrence of periodic dips in the learning curve resulting from all *Ss* turning pages at the same time.

Design and procedure.—All *Ss* received 20 30-sec. trials on the learning task. A massed practice group (MP) of 33 *Ss* received the 20 trials consecutively without a rest. A distributed practice group (DP) of 36 *Ss* was given a 30-sec. rest period between each trial. A third group of 42 *Ss* also received 30-sec. intertrial intervals, but instead of resting these *Ss* drew zeros during the interval with the same hand as that used for the learning task. This was called the same-hand group (SH). The fourth group of 42 *Ss* had conditions identical to those of the SH group except that these *Ss* printed zeros with the opposite hand from that used for the learning task. This was called the OH group.

Following Trial 20 all groups had a 10-min. rest period, and were then given three 30-sec. learning-task trials under massed-practice conditions.

RESULTS[3]

Figure 1 shows the performance curves for the four groups. In order

[3] A table giving individual trial-by-trial data for all *Ss* has been deposited, along with the verbatim instructions, with the American Documentation Institute. Order Document No. 4800, remitting $1.25 for 35-mm. microfilm or $1.25 for 6 × 8-in. photocopies.

TABLE 1

ANALYSIS OF VARIANCE OF PERFORMANCE ON TRIALS 2, 6, 11, 16, AND 19

Source	*df*	Mean Square	*F*	*P*
Trials	4	1612.48*	180.37	.001
Groups	3	1605.62	11.27	.001
T × G	12	71.25*	7.97	.001
Ss	149	142.46		
T × *Ss*	596	8.94		
Total	764			

* Tested by T × *Ss*.

to determine whether or not the experimental conditions produced significant differences in the means of the various groups, the scores on Trial 20 were subjected to an analysis of variance. The *F* ratio obtained indicated that the experimental conditions had a significant differentiating effect (*F* of 21.82, 3 and 149 *df*, $P < .001$).

As a basis for examining the difference between particular curves, the data for all four groups on Trials 2, 6, 11, 16, and 19 were subjected to an analysis of variance. These points were arbitrarily selected as representative of performance over all 20 trials.[4] A summary of the analysis appears in Table 1, and indicates highly significant slope as well as mean differences among the four curves.

The test of the prediction of depressed performance as a result of constant use of the same hand was made by comparing the prerest performance curves of groups MP and SH (see Table 2). The results of the analysis of variance of the scores of these groups on Trials 2, 6, 11, 16, and 19 showed a slope difference in

[4] A previous selection of Trials 1, 5, 10, 15, and 20 was abandoned because Bartlett's test revealed that the variances of the groups were heterogenous for these trials.

favor of group SH which was highly significant. This finding provides evidence against the prediction and, hence, against the effector-localization hypothesis.

The prediction of facilitation of performance through opposite-hand practice was tested by comparing the OH and DP groups (see Table 2). An analysis of variance of the scores of these two groups on Trials 2, 6, 11, 16, and 19 revealed the presence of a highly significant slope difference in favor of the DP group, thereby contradicting the prediction and providing more concrete evidence against the effector-localization hypothesis.

Another bit of evidence against the effector-localization hypothesis was the finding of no significant slope difference between the SH and OH groups ($F < 1.00$), suggesting that work inhibition is bilaterally transferable (see Table 2).

TABLE 2

ANALYSIS OF VARIANCE, TRIALS 2, 6, 11, 16, 19; PAIRS OF GROUPS

Source of Variation	*df*	Mean Square	*F*
F Test of groups MP-SH			
Trials	4	403.86	51.38
Groups	1	479.72	3.61
T × G	4	33.00*	4.20**
Ss	73	132.86	18.18
T × *Ss*	292	7.86	
Total	374		
F Test of groups SH-OH			
Trials	4	853.65	95.38
Groups	1	311.28	1.92
T × G	4	6.08*	<1.00
Ss	82	162.03	19.22
T × *Ss*	328	8.95	
Total	419		
F Test of groups OH-DP			
Trials	4	1347.29	135.00
Groups	1	981.69	6.47
T × G	4	42.08*	4.22**
Ss	76	151.66	15.19
T × *Ss*	304	9.98	
Total	389		

* Tested by T × *Ss*.
** $P = .001$.

TABLE 3

MEAN GAIN (LETTERS PRINTED) OVER 10-MIN. REST

Group	*N*	Gain	*t*	*P*
MP	33	3.3	4.78	.001
SH	42	3.7	6.17	.001
OH	42	4.5	4.34	.001
DP	36	1.4	2.37	.05

The *t*-test results for each of the groups, comparing Trial 20 with Trial 21, showed that the three groups OH, SH, and MP had highly significant amounts of reminiscence (see Table 3). The DP group did not show as much reminiscence, but did reach the .05 level.

An analysis of variance concerning the amount of reminiscence was run, comparing groups OH, SH, and MP. The *F* value was less than one, indicating that there was no significant difference in reminiscence for these groups.

DISCUSSION

The evidence against the predictions of facilitation of performance through opposite-hand practice and of depression of performance through same-hand practice on a different task[5] raises considerable doubt as to the tenability of a strict peripheral hypothesis for work inhibition. The presence of bilateral transfer of work inhibition also supports the findings of Kimble (**7**), who used a pursuit-rotor task. The necessity of assuming some sort of central factor as a supplement to the peripheral factor in work inhibition

[5] While it may be argued that not exactly the same muscle groups were being used in the two tasks, it is no doubt true that the differences were insignificant compared to the similarities.

is thus indicated. Two such central factors are immediately suggested.

The facilitating effect of the same-hand zero-drawing relative to massed practice could be due to the "relaxing" effect of this extremely simple interpolated activity. That is to say, some inhibition may be dissipated during intertrial intervals simply because *S*s of the SH group are not required to concentrate to the same degree as that required by the learning task. The depression of the OH performance function below distributed practice could also be explained in this way if one considers the large amount of concentration required by the performance of even this simple alternate task with the nonpreferred hand. While it may be true that this explanation could also apply in terms of muscular considerations, this does not prevent it from being considered as a possible central factor as well.

A second possible explanation of the results obtained is to assume a "monotony" or "boredom" factor. Monotony may be a performance depressor. It is possible that the major reason the SH group was significantly greater in slope performance than the MP group is that the MP group had conditions more conducive to boredom than did the SH group. The same reasoning could apply to the significant difference between the OH and DP groups. It is interesting to note that such a central factor is not contradictory to the obtained nonsignificant difference between the OH and SH groups. Of all the groups, these two were the most alike in their conditions of tasks and time intervals.

The fact that reminiscence occurred in the OH group is apparently contradictory to the peripheral hypothesis. According to this hypothesis, only the MP and SH groups should have reminiscence. The other two groups were presumably able to dissipate inhibition during intertrial rest periods. One possible interpretation of the highly significant postrest gain of the OH group is that this gain was the result of rehearsal during the 10-min. rest period. But this possibility is not exclusive to the OH group. In general, however, a rehearsal theory of reminiscence can not be entirely eliminated. At any rate, with the OH group showing as much reminiscence as the SH and MP groups, one is almost forced to accept the occurrence of bilateral transfer. This assumption is reinforced by the fact that the OH and SH groups showed no significant mean performance difference in their prerest trials.

Summary

Hull's effector-localization hypothesis of reactive inhibition (I_R) was tested, using the learning task of inverted alphabet printing and 169 introductory psychology students as *S*s.

Two groups were given 20 30-sec. trials on the learning task, together with 20 30-sec. intertrial periods devoted to performance of an alternate task: one group (SH) worked on the alternate task with the same hand as that used for the learning task, whereas the other group (OH) used the opposite hand for the alternate task work. The performance curves of these groups were compared with those of two additional groups, working under the usual massed- (MP) and distributed-practice (DP) conditions. The predictions tested were that opposite-hand practice would facilitate performance and that constant muscular work (same-hand) on different task would depress performance similar to the depressed performance of the massed practice group.

The results provide statistically significant evidence against both predictions and suggest, in addition, that bilateral transfer of work inhibition has taken place. These findings were interpreted as suggestive of a central inhibitory factor in addition to a muscular-fatigue type of factor in work inhibition.

References

1. Ammons, R. B. Effects of pre-practice activities on rotary pursuit performance. *J. exp. Psychol.*, 1951, **41**, 187–191.
2. Archer, E. J. Postrest performance in motor learning as a function of prerest distribution of practice. *J. exp. Psychol.*, 1954, **47**, 47–51.
3. Grice, G. R., & Reynolds, B. Effects of varying amount of rest on conventional and bilateral transfer "reminiscence." *J. exp. Psychol.*, 1952, **44**, 247–252.

4. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
5. Kientzle, M. J. Properties of learning curves under varied distribution of practice. *J. exp. Psychol.*, 1946, **36,** 187–211.
6. Kimble, G. A. Performance and reminiscence in motor learning as a function of degree of distribution of practice. *J. exp. Psychol.*, 1949, **39,** 500–510.
7. Kimble, G. A. Transfer of work inhibition in motor learning. *J. exp. Psychol.*, 1952, **43,** 391–392.
8. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
9. Mowrer, O. H., & Jones, H. M. Extinction and behavior variability as functions of effortfulness of task. *J. exp. Psychol.*, 1943, **33,** 369–386.
10. Rockway, M. R. Bilateral reminiscence in pursuit-rotor learning as a function of amount of first-hand practice and length of rest. *J. exp. Psychol.*, 1953, **46,** 337–344.
11. Schucker, R. E., Stevens, L. B., & Ellis, D. S. A retest for conditioned inhibition in the alphabet-printing task. *J. exp. Psychol.*, 1953, **46,** 97–102.
12. Siegel, P. S. Reactive inhibition as a function of number of response evocations. *J. exp. Psychol.*, 1950, **40,** 604–608.
13. Solomon, R. L. The influence of work on behavior. *Psychol. Bull.*, 1948, **45,** 1–40.
14. Wasserman, H. N. The effect of motivation and amount of pre-practice upon inhibitory potential in motor learning. *J. exp. Psychol.*, 1951, **42,** 162–172.

(Received May 13, 1955)

Journal of Experimental Psychology
Vol. 51, No. 5, 1956

RETENTION OF CONCEPTS AS A FUNCTION OF THE DEGREE OF ORIGINAL AND INTERPOLATED LEARNING[1]

JACK RICHARDSON

Northwestern University

It is usually taken for granted that materials such as prose, poetry, and concepts are retained better than less meaningful material, but this assumption seems to have attracted relatively little experimental attention. McGeoch states: "The higher retention values of meaningful materials are to be expected from the relatively low susceptibility of these materials to interference from interpolated learning" (**4**, p. 367). Osgood, on the other hand, observing that we would infrequently encounter nonsense syllables in everyday experience, maintains: "In fact, if it were clearly demonstrated that nonsense syllables were less well retained than meaningful words, it would constitute presumptive evidence against the theory of forgetting which McGeoch held" (**9**, p. 562).

No investigator has clearly demonstrated the direct relation between retention and meaningfulness for rote-learned materials (**16**). Concepts are usually considered to be more meaningful than rote material. Reed (**10, 11, 12, 13**) found retention of concepts after 1, 2, 3, and 6 weeks to be high and infers that it is higher than rote-learned material. There are several methodological problems involved, however, which make such an inference questionable, and there is no firm basis for concluding that the difference in retention for rote materials and concepts is to be attributed to meaningfulness. The present study, while not dealing with meaningfulness and retention directly, is an attempt to study retention of concepts as a function of variables which have classically produced large variation in retention of rote-learned materials.

The retroactive inhibition (RI) paradigm has been used many times in verbal learning to demonstrate the effect of the amount and the type of interpolated learning upon the retention of previously learned material but, as far as we can determine, this paradigm has never been employed with concept materials. McGeoch and McKinney found small amounts of RI using both prose (**6**) and poetry (**5**). Sisson (**14**) found that nonsense syllables of lower association value were less subject to interference from interpolated learning than those of higher association value. This seems to be the only study which directly compares the effect of interpolated learning on the retention of materials which differ in meaningfulness, and the results are opposite to those predicted by McGeoch.

Underwood and Richardson (**17**) have developed concept material which is flexible enough to be used in the RI paradigm. This material consists of concrete nouns and adjectives which describe the objects named by the nouns. The nouns were presented to *S*s who were asked to give one immediate sensory term which described the object named by the noun. The descriptive adjectives given as responses to the nouns are used as the concept names and the nouns are used as the examples of the concepts. "Dominance" is defined by the percentage of *S*s who gave a particular adjective as the response to a noun with the higher dominance level corresponding to the greater percentage

[1] This study is abstracted from a dissertation submitted to the Graduate School, Northwestern University. The author is indebted to Dr. Benton J. Underwood for his supervision.

of the particular response. Thus if *white* were 80, 71, 50, and 37% of the responses given to *chalk*, *snow*, *diaper*, and *pearl*, respectively, then these percentages define the dominance of the concept *white* when applied to each of the examples. "Overlap" is defined as the number of times that an example of one concept could also be an example of another concept. If *apple* were described as both red and round and is used as an example of *red* in a group of concepts where *round* also appears, then one of the examples of *red* could also be an example of *round* and there is an overlap of the two concepts. Preliminary investigation demonstrated that *S*s who learned a list of low-dominance, minimum-overlap concepts to a relatively low degree showed no forgetting after a 7-day retention interval.

In verbal learning it has been found that increasing the degree of original learning increases retention (**2**) and increasing the degree of interpolated learning decreases retention (**1, 3, 7, 15**). This study is designed to determine whether the two variables have a similar effect on the retention of concepts.

Method

Materials.—Two lists were constructed from the materials described above in which the stimuli (nouns) were the same but the concepts (adjectives) designated as correct were different. Thus, *bone*, for which 34% of the responses were *white* and 47% of the responses were *hard*, is used as an example of *white* in one list and as an example of *hard* in the other. Each list consists of the same 16 nouns: four examples of each of four concepts. Of the four examples of any one concept in one list, two are examples of the same concept in the other list, while the other two are examples of two different concepts. Thus each concept in one list overlaps three concepts in the other list. The intralist overlap was minimized. Owing to the limited number of rated words, it was necessary to use one example which overlapped with another concept within the list. The stimuli presented to *S* were identical for the two lists and the correct concepts could be discovered only by giving responses and being told "right" or "wrong" by *E*. This will be recognized as the A–B, A–C paradigm of verbal learning except that the correct responses were not presented to *S* and each correct response was correct for four of the stimuli.

TABLE 1

Concept Examples and Concept Names

List 1	
"White"	"Shines"
Bone	Bracelet
Hospital	Armor
Goat	Diamond
Bread	Silk
"Small"	"Black"
Pill	Derby
Sardine	Pot
Kitten	Asphalt
Cradle	Skunk
List 2	
"Round"	"Hard"
Pill	Bone
Bracelet	Armor
Derby	Diamond
Pot	Asphalt
"Smells"	"Soft"
Hospital	Bread
Goat	Kitten
Sardine	Cradle
Skunk	Silk

The mean dominance levels of the two lists was 29.5% and 31.4%. The 16 examples of the concepts were typed on white vellum tape in three different orders. The examples were assigned randomly to a position within an order subject to the restriction that two examples of the same concept did not appear in succession. Only the examples of the concepts appeared on the tape and *S*s were not told and did not see the concept names at any time.

Design.—The usual RI design—Learn List 1—Learn List 2—Relearn List 1—was used. The original learning (OL) was to a criterion of either 8, 12, or 16 correct responses on a single trial and the interpolated learning (IL) to a criterion of either 0, 8, or 16 correct responses on a single trial. The experimental design may be diagrammed as follows:

		OL Criterion		
		8	12	16
IL Criterion	0	10 *S*s	10 *S*s	10 *S*s
	8	10 *S*s	10 *S*s	10 *S*s
	16	10 *S*s	10 *S*s	10 *S*s

All relearning occurred approximately 24 hr. after the OL and was carried to the same

criterion as OL. The two lists were counterbalanced so that each appeared equally often for all degrees of OL and IL. The numbers used to refer to the nine groups will designate the degree of OL or degree of IL by giving the learning criterion. Thus the 8-OL groups will refer to those that learned the original list to a criterion of eight correct responses on a single trial. Each *S* served in only one condition.

Procedure.—Each *S* received instructions to respond to each example and to limit his responses to the immediate sensory level. The nouns were presented on a Hull-type memory drum at a 3-sec. rate with a 3-sec. intertrial interval. The *S* was permitted to respond any time during the 3-sec. interval and was told "right" or "wrong" after each response. There was approximately 1 min. between the OL and IL, and *S* was told that the correct responses for the IL list would be different although the nouns would be the same. After IL *S* was cautioned not to mention the experiment to anyone. The *S* returned the next day at the same time and relearned the first list.

Results

Original and interpolated learning.—The mean number of trials of OL required to learn to a criterion of 8, 12, and 16 correct responses was 6.1, 10.7, and 17.2, respectively. The mean number of trials of IL required to learn to a criterion of 8 and 16 correct responses was 8.9 and 15.5, respectively. As might be expected, these differences due to the learning criteria are highly significant.

Many *S*s did not learn to the assigned criterion during the 1-hr. experimental period so there is no pretense of random assignment to groups. The number of trials required to learn the first list to a criterion of eight correct responses was used to match the nine groups. A total of 56 *S*s was dropped from the experiment, 50 because they did not learn within the allotted time and 6 for miscellaneous reasons such as failure to understand the instructions or failure to return for relearning. Selection of *S*s could possibly result in groups which were biased with respect to IL or retention. The product-moment correlation between trials to learn to a criterion of 8 correct on OL and trials to learn to the same criterion on IL was $-.09$. This is not significant for an N of 60, so performance on IL cannot be predicted from performance on OL.

Most *S*s dropped from the experiment were slower than those retained and it is possible that speed of learning is related to retention. As a check on this, *S*s within each group were divided into fast and slow learners on the basis of trials required to learn the first list. This resulted in two groups of 45 *S*s each and the mean number of trials required for OL was 16.18 and 6.24. The mean number of correct responses on the first relearning trial was 10.96 and 11.04; the mean number of items lost (the number correct on the last trial of OL minus the number correct at recall) was 1.44 and 1.29 for the fast and slow groups, respectively. The differences between these means do not approach significance, so evidently the selection of *S*s did not bias retention.

Relearning.—Owing to marked heterogeneity and lack of knowledge about the distributions of the relearning and retention scores the median test (**8,** p. 404) was used to evaluate the significance of the differences. The median number of trials required for each group to relearn the original list after 24 hr. is presented in Table 2. Degree of OL did not affect the trials to relearn ($\chi^2 = .26$) but the differences due to degree of IL are significant above the 1% level ($\chi^2 = 13.15$). The number of trials required to relearn increase as the degree of IL increases.

Loss scores.—As a measure of forgetting over the 24-hr. interval, a loss score was calculated for each *S* by subtracting the number of correct

TABLE 2

MEDIAN NUMBER OF TRIALS REQUIRED TO RELEARN

		OL					
		8		12		16	
		Median	Range	Median	Range	Median	Range
IL	0	1.0	1–1	1.2	1–3	1.7	1–3
	8	1.7	1–6	3.0	1–16	2.0	1–3
	16	3.7	1–7	3.5	2–13	2.0	1–5

responses on recall from the number of correct responses on the last trial of OL. This loss score is a measure of the items lost during the interval, but should be corrected for learning on the last trial of OL if it is to be uncontaminated by degree of learning. There is, within the experiment, a control for the groups which learned to a lower criterion than one perfect trial. If the groups which learned the original list to one perfect trial are used to calculate the gain on the trials after they reached the other criteria, i.e., on the trial after 8 correct responses and the trial after 12 correct responses, then this is an estimate of the learning on the last trial of OL for the groups which learned to a criterion of 8 correct or to a criterion of 12 correct responses. Actually, the groups which learned to a criterion of 12 correct responses were also used in computing the gain after the trial on which the criterion of 8 correct responses was attained. The mean gain on the trial after the criterion of 8 correct was .87 items and the mean gain on the trial after the criterion of 12 correct was .60 items. If this mean gain is added to each *S*'s score on the last trial of OL in the appropriate groups and if we assume that the learning during the trial on which the criterion of one perfect is attained has a negligible effect on recall, then by subtracting the items correct on recall from these we have an estimate of loss uncontaminated by differences in rate of learning at different points on the acquisition curve.

The medians of the corrected loss scores are plotted in Fig. 1. The differences due to the degree of OL are not significant ($\chi^2 = 4.64$), but the differences due to degree of IL are significant at the 1% level ($\chi^2 = 26.55$). Although the differences due to degree of OL are not significant, there appears to be an interaction between degree of OL and degree of IL. The 0-IL groups show a significant increase in loss as degree of OL increases ($\chi^2 = 12.37$), but the other degrees of IL first show an

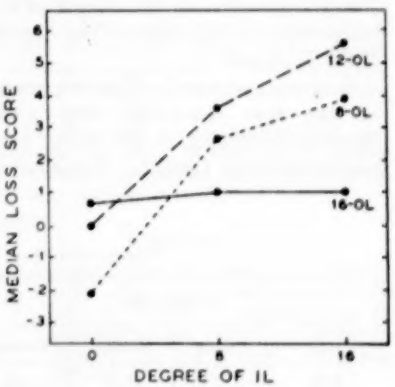

FIG. 1. The median number of items forgotten during the 24-hr. interval.

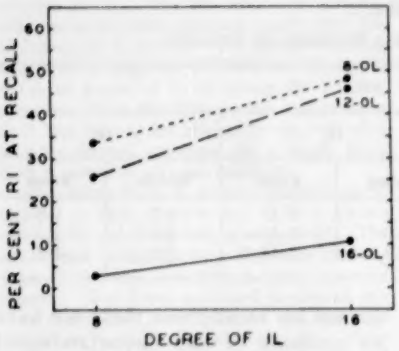

FIG. 2. Percentage of RI on the first relearning trial.

increase and then a decrease as degree of OL increases. These are not significant but the differences in the 16-IL groups almost reach significance ($\chi^2 = 5.41$; $P < .10$).

As the degree of IL increases, the loss over the retention interval increases, thus showing the same thing as the relearning scores; IL does interfere with retention and the greater the degree of IL the more interference. The percentage of RI for each group, computed on the number of correct responses at recall, is plotted in Fig. 2. The RI increases as the degree of IL increases, and decreases as the degree of OL increases.

Responses at recall.—The responses on the first relearning trial were classified according to the source and are presented in Table 3. The entries in the table are percentages of the total responses for all three groups which received the degree of IL indicated at the left. The column headings refer to the source of the response. "OL Responses" and "IL Responses" refer to the list from which the response came, and "Correct" and "Incorrect" indicate whether the response was correct for the noun to which it was given or whether it was incorrect but appropriate for other nouns in the designated list. "Omissions" refers to the failure of *Ss* to give any overt response while "Rated" are responses which were given by at least 5% of the *Ss* on the original rating of the material. "Other" includes all responses not falling into the preceding categories. Since *Ss* were asked to recall the first list, the first column gives the percentage of correct responses and the other columns the percentages of errors of various types.

The responses from the original list decrease as IL increases while all other categories, except "Other," increase. However, most of the variation due to degree of IL is accounted for by the variation in the direct intrusions from the interpolated list and by failure of *Ss* to respond. If the percentages of responses that were correct for either OL or IL and the percentage of times that *Ss* did not give a response are summed, these three categories account for 83.54, 83.13, and 85.63% of the responses in the groups which

TABLE 3

CLASSIFICATION OF RESPONSES AT RECALL

Groups	OL Responses		IL Responses		Omissions	Rated	Other
	Correct	Incorrect	Correct	Incorrect			
0-IL	83.33	11.67	.62	.62	1.67	.42	1.67
8-IL	67.29	10 42	9.37	1.67	6.46	1.46	3.33
16-IL	54.58	8.33	19.17	2.08	9.79	3.12	2.92

learned the interpolated list to a criterion of 16, 8, and 0 correct responses on a single trial, respectively.

Discussion

Using the A–B, A–C paradigm it has been shown that retention of concepts decreases as the degree of IL increases. In verbal learning studies it has been demonstrated that decrement in retention is a negatively accelerated increasing function of the number of trials of IL, and that the number of intrusions from the interpolated list decrease with the higher degrees of IL (**2, 7, 15**). The loss scores presented in Fig. 1 give little indication that the function is a negatively accelerated one, and the intrusions, presented in Table 3, increase as the degree of IL increases. However, these apparent differences may well be due to differences in degree of learning.

Recall of the concepts by the 8- and 12-OL groups which had no IL shows a gain over the last trial of learning. The 16-OL group which did not receive IL had a mean recall score of 15.4 items. Apparently this is much better retention than that found in verbal learning but again the two tasks are not equated on potentially important variables. The differences in retention may be due to the small number of concepts learned, the degree of learning, or to associating several stimuli with the same response in the concept task.

It is possible that the loss due to IL may be the result of the loss of concept names, loss of specific examples, or lack of gain of correct examples shown by the 0-IL groups. A detailed analysis of the last learning trial and the first relearning trial shows that all three vary with the degree of IL. However, most of the decrement is due to loss of specific examples of concepts and not to the loss of concept names. Only 25 names were forgotten by all 90 *Ss*; 18 of these were in the 16-IL groups and 7 in the 8-IL groups.

From this study it can only be concluded that concepts, as defined, are subject to interference from IL and that the interference increases as IL increases. The interference is primarily with specific examples and seldom results in complete loss of a concept.

Summary

The usual RI paradigm, with A–B, A–C lists of concepts was used to study the effect of degree of OL and degree of IL on retention. The material consisted of concrete nouns as examples of the concepts and descriptive adjectives as the concept names. The correct concept names were not presented at any time and the *Ss* gave responses until they discovered the correct name. A factorial design was used with the original list learned to a criterion of either 8, 12, or 16 correct responses on a single trial and the interpolated list learned to a criterion of either 0, 8, or 16 correct responses on a single trial. After a 24-hr. retention interval, *Ss* relearned the original list.

The results may be summarized as follows: (*a*) Retention decreased as degree of IL increased. (*b*) These differences in retention could be accounted for by the intrusions from the interpolated list and failures to respond. (*c*) The degree of OL was not an effective variable on retention but there appeared to be an interaction between degree of IL and degree of OL. Without IL there was a decrease in retention as OL increased while, with IL, the retention first decreased and then increased as degree of OL increased.

The two groups which learned the original list to a criterion less than one perfect trial and which did not receive IL gave more correct responses at recall than on the last trial of OL. This gain consisted of examples of concepts that had been given correctly during learning even though the particular example had not elicited the correct concept name. There was no forgetting of concept names in the groups which did not receive IL.

The decrement in retention due to IL was primarily the result of the loss of specific examples of concepts that were given correctly for other examples at recall.

References

1. Archer, E. J., & Underwood, B. J. Retroactive inhibition of verbal associations as multiple function of temporal point of interpolation and degree of interpolated learning. *J. exp. Psychol.*, 1951, **42**, 283–290.
2. McGeoch, J. A. The influence of degree of learning upon retroactive inhibition. *Amer. J. Psychol.*, 1929, **41**, 252–262.

3. McGeoch, J. A. The influence of degree of interpolated learning upon retroactive inhibition. *Amer. J. Psychol.*, 1932, **44,** 695–708.

4. McGeoch, J. A. *The psychology of human learning.* New York: Longmans, Green, 1942.

5. McGeoch, J. A., & McKinney, F. Retroactive inhibition in the learning of poetry. *Amer. J. Psychol.*, 1934, **46,** 19–33.

6. McGeoch, J. A., & McKinney, F. The susceptibility of prose to retroactive inhibition. *Amer. J. Psychol.*, 1934, **46,** 429–436.

7. Melton, A. W., & Irwin, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53,** 173–203.

8. Mood, A. McF. *Introduction to the theory of statistics.* New York: McGraw-Hill, 1950.

9. Osgood, C. E. *Method and theory in experimental psychology.* New York: Oxford Univ. Press, 1953.

10. Reed, H. B. Factors influencing the learning and retention of concepts: I. The influence of set. *J. exp. Psychol.*, 1946, **36,** 71–87.

11. Reed, H. B. The learning and retention of concepts: II. The influence of length of series: III. The origin of concepts. *J. exp. Psychol.*, 1946, **36,** 166–179.

12. Reed, H. B. The learning and retention of concepts: IV. The influence of the complexity of the stimuli. *J. exp. Psychol.*, 1946, **36,** 252–261.

13. Reed, H. B. The learning and retention of concepts: V. The influence of form of presentation. *J. exp. Psychol.*, 1950, **40,** 504–511.

14. Sisson, E. D. Retroactive inhibition: the influence of degree of associative value of original and interpolated lists. *J. exp. Psychol.*, 1938, **22,** 573–580.

15. Thune, L. E., & Underwood, B. J. Retroactive inhibition as a function of degree of interpolated learning. *J. exp. Psychol.*, 1943, **32,** 185–200.

16. Underwood, B. J. Speed of learning and amount retained: a consideration of methodology. *Psychol. Bull.*, 1954, **51,** 276–282.

17. Underwood, B. J., & Richardson, J. Some verbal materials for the study of concept formation. *Psychol. Bull.*, 1956, **53,** 84–95.

(Received June 6, 1955)